PAYOFFS TO PROBABILITY FORECASTERS*

by

Arlo D. Hendrickson

Technical Report No. ~~144~~ 145

University of Minnesota
Minneapolis, Minnesota

September 1970

*Part of this material is taken from two earlier reports. Chapters II and III are based in part on Technical Report No. 141, "Payoffs of probability forecasters and a theorem of McCarthy," by Arlo D. Hendrickson (1970); Chapter IV is based in part on Technical Report No. 133, "Some procedures for determining subjective probabilities by sequential choices," by Arlo D. Hendrickson and Robert J. Buehler (1970). This work was supported by National Science Foundation Grant GP-9556.

Arlo D. Hendrickson
435 words

ABSTRACT

Payoffs to Probability Forecasters

A payoff $f(p)$ is defined to be a function whose domain consists of densities with respect to a measure μ on a measure space $(\mathfrak{X}, \mathfrak{a})$ and whose range is a class of random variables on $(\mathfrak{X}, \mathfrak{a})$. The value of a payoff $f(\hat{p})$ after the outcome $\omega \in \mathfrak{X}$ is observed is the payment given to a forecaster for his appraisal $\hat{p}$ of the true density p. A desirable property of a payoff function is that the maximum expectation be attained under a given density p when $\hat{p} = p$. This property is said to "keep the forecaster honest."

A theorem of McCarthy (1956) on payoff functions which keep the forecaster honest is corrected by a slight modification and is generalized to a Hilbert space $\mathcal{H}$, which is taken to be a closed subspace of $\mathcal{L}_2(\mu)$. If both the domain $\mathcal{P}$ and the range of f are subsets of $\mathcal{H}$, then the expectation $H(p) = E(f(p)|p)$ is the usual inner product $< p, f(p) >$. By generalizing Rockafellar's finite dimensional definition of subgradient, the condition for encouraging honesty becomes equivalent to the condition that the expectation function $H(p)$ have a subgradient $f(p)$ at each point $p \in \mathcal{P}$. The class of continuous, convex and homogeneous functions on the cone $\{\lambda p: p \in \mathcal{P}, \lambda > 0\}$ are among the class of expectation functions of payoff rules which encourage honesty.

Payoff functions are also studied from the point of view of offering a forecaster a <u>choice</u> of several random variables without

asking him to disclose his evaluation of the density. The class C of choices is assumed to be closed and convex. There is a one-to-one correspondence between closed convex sets C and expectation functions H of honesty encouraging payoff rules f. H is characterized as being a support function of C where p is normal to C at $f(p)$ for all $p \in \mathcal{P}$.

A sequential procedure is developed to obtain the forecaster's undisclosed probability $p = P(A)$. Let $r_1 = \frac{1}{2}$. At the $k\underline{th}$ stage the forecaster is offered his choice between the payoff $g(p_k)$ if an event B_k of probability p_k occurs. If the payoff when A occurs is chosen then $p_{k+1} = p_k + 2^{-k-1}$. Otherwise $p_{k+1} = p_k - 2^{-k-1}$. It is assumed that the forecaster makes his decision such that $p \in (p_k - 2^{-k}, p_k + 2^{-k})$ for all k. To ensure this we let $\hat{p} = \lim_{k \to \infty} p_k$ and ask that g be such that the forecaster's cumulated payoff $f(\hat{p})$ encourages honesty. Necessary and sufficient conditions on g are given to encourage honesty at either time n or in the limit as $n \to \infty$. One solution is given by $g(2^{-n}) = 2^{-n}$ and $g(\frac{k}{2^n}) = \frac{2 \cdot 4 \cdot \ldots (k-3)(k-1)}{3 \cdot 5 \cdot \ldots (k-2)k} \; 2^{-n}$ for $k = 3, 5, \ldots, 2^{-n} - 1$, and $n = 1, 2, \ldots$.

To

My Wife

Sonya

ACKNOWLEDGEMENTS

It is a pleasure to acknowledge the invaluable help provided by others.

I owe many thanks to Professor R. J. Buehler under whose excellent guidance this thesis was written. Professor Buehler provided much of the literature on payoffs and suggested the problems of generalizing McCarthy's Theorem and of developing an honesty encouraging sequential procedure. Professor Buehler offered much encouragement and gave many suggestions which considerably improved this thesis.

I would like to thank the National Science Foundation for their financial assistance under the research grant GP-9556. I would also like to thank the U. S. Bureau of Mines, Twin Cities Mining Research Center with whom I was employed part time while completing research on this thesis.

I am grateful to Mr. Gerry DuChaine who typed this thesis. Mr. DuChaine's ability as a math typist made proofreading easy and expedited the final typed draft of this thesis.

Finally I extend warm thanks to the Statistics Faculty of the University of Minnesota under whom I spent many fulfilling hours as a student and who contributed indirectly to this thesis.

TABLE OF CONTENTS

CHAPTER I

Introduction

1. Background.

Suppose a forecaster has knowledge of a probability measure, unknown to others, for a given measure space. It is not essential to this paper whether the forecaster's probability is subjective or is an estimate of a unique probability measure defined by a frequency interpretation. A forecaster's client would want the forecaster to be honest when giving his subjective probability, or to be accurate when assessing the true probability. The client can encourage honesty or provide incentive for accuracy by giving the forecaster a payoff which depends on both the forecaster's disclosed probability and on the outcome of the experiment $\omega \in \mathfrak{X}$. The class of such "payoffs which encourage honesty" will be studied in Chapter II.

Related to payoff functions which encourage honesty is the concept of score functions for forecasters which encourage accurate prediction, first suggested by Brier (1950). If $E_1, E_2, \ldots, E_r$ are r disjoint, exhaustive outcomes of any given trial of an experiment, and if p_{ij} is the probability given by a forecaster for the event E_i at the jth trial, then Brier suggested the score

$$P = n^{-1} \sum_{j=1}^{n} \sum_{i=1}^{r} (p_{ij} - \gamma_{ij})^2$$

where γ_{ij} is the indicator function for the event E_i at the jth trial, and n is the number of trials. The score P is minimized when the prediction is "perfect," when $p_{ij} = \gamma_{ij}$. If we take $n = 1$, negate P, and add the constant 1, we obtain the quadratic payoff, $2p_k - \sum_{i=1}^{r} p_i^2$ if the event E_k occurs. This payoff function, suggested by de Finetti (1962) is known to satisfy the property of encouraging honesty, which we will define below. Brier did not consider the expectation of his score.

Good (1952), in a section titled "fair fees," raised the question of "how a firm can encourage its experts to give fair estimates of probabilities." If E was an event whose probability was estimated to be p_1 by an expert, then Good suggested paying the expert $k \log (2p_1)$ if E occurred and $k \log (2-2p_1)$ if E did not occur. This payoff function had the "desirable property that its expectation is maximized when $p_1 = p$, the true probability of E." Good (1970) has stated that a more appropriate term for "fair fees" would have been "accuracy incentive fees."

McCarthy (1956) generalized Good's problem from the case of two possible outcomes to that of n outcomes whose probabilities were to be estimated by a forecaster. A forecaster was to be paid a payoff $f_k(q)$ if the kth event occurred where $q = (q_1, q_2, \ldots, q_n)$ was the forecaster's estimate of the true probability vector $p = (p_1, p_2, \ldots, p_n)$. McCarthy defined a payoff rule which "keeps the forecaster honest" to be a rule such that "regardless of the value of p, the forecaster's expectation is maximized if and only if he puts $q = p$." In this paper we will call such functions "payoff functions which strictly encourage honesty." We reserve the name "payoff functions which encourage honesty" for payoff functions which satisfy the less restrictive condition that the forecaster's expectation is maximized at $q = p$ and possibly other values of q. At least this condition does not discourage honesty, although it may not provide incentive for collecting more evidence.

Several authors have studied the properties of payoff functions which encourage honesty. Aczel and Pfanzagl (1965) obtained the interesting result that with $n > 2$ outcomes, the logarithmic payoff, $A \log p_k + B$, is the only function which both strictly encourages honesty and has the property that the payoff for the occurrence of the kth event depends only on the estimated probability p_k of that event. They give more general solutions for $n = 2$. For other more or less independent discussions of this same result see McCarthy (1956), Marschak (1960), Shuford et al (1966), de Finetti and Savage (1969), and Good (1970).

Denote by $H(r, p)$ the expected payoff to an appraiser for his appriasal of the true probability vector p as being r. De Finetti and Savage (1969) have shown the condition that the loss function $L(r, p) = H(p, p) - H(r, p)$ be dependent only on $r - p$ is equivalent to the condition that $H(p, p)$ be quadratic, and this is equivalent to the condition that $L(r; p) = L(p; r)$.

The most general theorem on payoffs which encourage honesty seems to be McCarthy's (1956) in which he gives necessary and sufficient conditions in the finite discrete case. He omitted the proof. McCarthy's theorem can be interpreted as follows: A random variable $f(p) = (f_1(p), f_2(p), \ldots, f_n(p))$ keeps the forecaster honest iff $f_k(p) = \frac{\partial H}{\partial p_k}(p)$ where H is a convex function which is homogeneous of the first degree. The function H is the maximum expectation function

$$H(p) = \sum_{k=1}^{n} p_k f_k(p).$$

McCarthy's theorem needs a slight modification to be correct, as we show in Chapter II.

Buehler (1970) gave examples of payoff functions where the domain was the set of density functions with respect to Lebesgue measure. Buehler posed the problem of paralleling the results of McCarthy's theorem for the Euclidean case with a theorem for the space of density functions.

Marschak (1960) in his comments on McCarthy's paper noted the distinction between functions which are "expected costs to the client" to provide incentive for honest appraisals by the expert and functions which are "a good measure of worth to the client to be given these probabilities." However, although the two measures may be different, according to McCarthy's theorem the former is restricted only to be convex (because any function which is convex on $\{(p_1, p_2, \ldots, p_n): \sum_{k=1}^{n} p_k = 1, p_k \geq 0$ for all $k\}$ can be extended uniquely to a homogeneous and convex function on the Euclidean space R^n). Thus if the second measure is restricted to be convex, then by an appropriate choice of the payoff function, the two measures of worth and cost can be made equal.

We quote from McCarthy (1956) the intuitive concept of the convexity restriction on the maximum expected payoff function H:

> The intuitive concept of the convexity restriction is that it is always a good idea to look at the outcome of an experiment if it is free. For suppose that the experiment has two outcomes, A and A^*, which would give one probabilities p and p^* for the event in question. Let t be the probability that A is the outcome. If we decide not to look, our expectation is $H(tp + (1-t)p^*)$, while if we decide to look, our expectation is $tH(p) + (1-t)H(p^*)$.

The payoff to a forecaster for his evaluation of p can be studied from a different viewpoint. Rather than paying a forecaster for disclosing a probability density, a client can offer the forecaster a choice of several random variables whose values depend on the outcome of the experiment $\omega \in \mathfrak{X}$. The forecaster's choice should yield information about his undisclosed evaluation of p. Several choices can be offered sequentially. The set of choices at each stage may depend on previous choices, and each particular choice may yield more information about the forecaster's probability density.

It will be assumed that the value of a payoff to a forecaster is in utility and not in money; because money usually has a diminishing marginal utility. If the payoff is positive and bounded by M, this can be achieved if one has a commodity whose utility is M and a device for randomizing with any size probability $0 \leq p \leq 1$. In the sequential case, this solution does not apply because utility is not additive. This is solved not by paying increments in utility at each stage, but instead by paying one cumulated utility. For a philosophical study of utilities or "desirabilities" and subjective probabilities or "degrees of belief" see for example Ramsey (1965) or Jeffrey (1965).

2. Summary.

Chapters II and III apply the theory of convex analysis to studying the properties of a payoff function f and its corresponding expected payoff function H. Theorem 7 of Chapter II generalizes McCarthy's theorem and includes Buehler's (1970) class of payoffs defined on the domain of continuous distributions. This generalization is accomplished by generalizing Rockafellar's (1970) definition of subgradient from Euclidean space to Hilbert space. Theorems 8, 9, and 10 give additional conditions on H to ensure that there exists f satisfying the requirements of Theorem 7.

Chapters III and IV study payoff functions from the viewpoint of presenting a forecaster a choice of several random variables, whose value is to be given to the forecaster after the outcome of the experiment is observed. If the forecaster is to maximize his expected payoff, we can assume without loss of generality that the class C of choices is convex. The class C is also taken to be closed, because it is desired that a payoff with maximum expectation over C be a member of C. Chapter III studies the realtionship between payoff functions and closed convex classes of random variables. A one-to-one dual correspondence is given between closed convex sets C of choices and functions H which are maximum expectations of honesty encouraging payoff functions f.

In Chapter IV, a method is described in which the forecaster is offered a sequence of choices, where each offer may depend on the previous choices. A naive person at each stage might make the "honest choice" in which he maximizes his increment in expected payoff. However, a person who has knowledge of the procedure might make a "dishonest choice" at a given stage in order to increase his cumulated expected payoff. We consider, in Chapter IV, the properties of procedures in which the consistently honest choices at each stage lead to a maximum cumulated expected payoff at time n or in the limit as $n \to \infty$. The problem is simplified by considering a choice at each stage between only two random variables whose expectations depend only on one fixed unknown probability $P(A)$.

The sequential procedure is described in detail in Chapter IV. Necessary and sufficient properties are given for the classes of increments in payoffs which encourage honesty at either the nth stage, at every stage, or in the limit.

Some problems are sketched in Chapter V which are not considered earlier in the paper. Comments are given on obtaining a consensus when a group of forecasters give several estimates of a density, on sequential procedures not considered in Chapter IV, on other desirable properties of payoffs besides that of encouraging honesty, and on generalizing McCarthy's theorem to undominated classes of probability measures.

CHAPTER II

Honesty Encouraging Payoff Functions and a Theorem of McCarthy

1. Introduction.

Payoff functions are rewards to forecasters for their disclosure of either their subjective probability measure or their assessment of a unique but unknown probability measure with a frequency interpretation. A payoff function depends on both the disclosed probability measure and the future outcome of the corresponding random experiment. As we have mentioned in Chapter I, several authors have considered the use of payoff functions for either defining subjective probabilities or evaluating the performance of probability appraisers.

McCarthy (1956) gave necessary and sufficient conditions on payoff functions to encourage honesty in the Euclidean case where the unknown probability vector is n-dimensional. McCarthy's theorem was stated somewhat ambiguously without proof. The theorem states that f is a payoff function which encourages honesty iff f is the gradient of a convex function H which is homogeneous of degree one. The function H satisfies $H(p) = E(f(p)|p)$ for all n-dimensional probability vectors p.

There has been some confusion about McCarthy's Theorem, even though it needs only a slight modification to be precise. For example, Marschak (1960, page 97) argues in a footnote that although the honesty encouraging logarithmic payoff, $f_k(p) = \log p_k$ where $p = (p_1, p_2, \ldots, p_n)$, is the partial derivative of a convex function

H, the function H can not be homogeneous. This is not true: every convex function H on a space $\mathcal{P}$ of probability vectors <u>can</u> be extended to a homogeneous function on $D = \{\lambda p : p \in \mathcal{P}, \lambda > 0\}$ simply by defining $H(\lambda p) = \lambda H(p)$. The function defined from the gradient of H in the logarithmic case is $f_k^*(p) = \log \frac{p_k}{\Sigma p_j}$ which encourages honesty not only on $\mathcal{P}$ but also on D. It seems natural to take all payoffs to be homogeneous of degree zero on D, and the corresponding expected payoff functions H to be homogeneous of degree one.

2. <u>Payoff functions</u>.

Let $(\mathfrak{X}, \mathfrak{A}, \mu)$ be a measure space and let $\mathcal{P}$ be a convex set of densities on $(\mathfrak{X}, \mathfrak{A})$ with respect to μ. Let $\mathcal{L}$ be the set of real-valued random variables on $(\mathfrak{X}, \mathfrak{A})$. We will define a "payoff function " to mean any function f which maps $\mathcal{P}$ into $\mathcal{L}$. Hence the value of $f(p)$ will be dependent on p and the outcome of the experiment, $\omega \in \mathfrak{X}$.

If a payoff is defined to be a real-valued random variable on $(\mathfrak{X}, \mathfrak{A})$, then the function $f(p_1)$ can be considered as the payoff given to an assessor for estimating p_1 as the true probability density. The assessor's choice $p_1 \in \mathcal{P}$ may actually be presented as an infinite sequence of choices and f may be an infinite series of increments in payoffs depending only on previous choices, as in Chapter IV. These increments in payoffs may arise from bets, they may be awards or penalties, and may have known distributions or unknown distributions depending on p.

Let us assume that a person, when given the choice among the random payoffs in the range of f, will select a random payoff which he believes to have maximum expectation. Then a payoff function will keep that person honest if when offered $f(p_1)$ for telling us p_1, he must tell us the true density p in order that he maximizes his expectation, $E(f(p_1)|p)$. Hence, we will say a payoff strictly encourages honesty if for $p, p_1 \in \mathcal{P}$ such that $p_1 \neq p$ a.e. μ,

(1) $$E(f(p)|p) > E(f(p_1)|p).$$

We will say f encourages honesty if f satisfies the less restrictive inequality

(2) $$E(f(p)|p) \geq E(f(p_1)|p) \quad \text{for all } p_1, p \in \mathcal{P}.$$

If (2) is satisfied, then in order to maximize his expectation with respect to p, an assessor must choose his assessment of p from the class

(3) $$B(p) = \{p_1 \colon E(f(p)|p) = E(f(p_1)|p)\}.$$

Thus, if an assessor chooses p_1 for the questioned density and is allowed to be dishonest, then (2) implies the assessed value of p is a member of

(4) $$A(p_1) = \{p: E(f(p)|p) = E(f(p_1)|p)\}.$$

If (2) holds, it will be shown in Section 6 that $A(p_1)$ is a convex set.

Example 1.

Let $P(E) = p$ be unknown and let $P(E_1) = \frac{1}{2}$. Define the payoff function f by

$$\text{if } p \geq \frac{1}{2} \text{ then } f(p) = \begin{cases} 1 & \text{if } E_c \text{ occurs} \\ 0 & \text{if } E^c \text{ occurs} \end{cases}$$

$$\text{if } p < \frac{1}{2} \text{ then } f(p) = \begin{cases} 1 & \text{if } E_1 \text{ occurs} \\ 0 & \text{if } E_1^c \text{ occurs.} \end{cases}$$

Then f satisfies condition (2). If $\mathcal{P} = [0, 1]$ then $A(\frac{1}{4}) = [0, \frac{1}{2}]$ and $A(\frac{3}{4}) = [\frac{1}{2}, 1]$. See Chapter IV.

Example 2.

Let $||p||^2 = \int p^2(x)d\mu(x) < \infty$ for all $p \in \mathcal{P}$. Define the random variable $f(p)$ by

$$f(p)(x) = \frac{p(x)}{||p||} \text{ for all } x \in \mathcal{X},\ p \in \mathcal{P}.$$

Then condition (1) is satisfied by the Cauchy-Schwarz inequality.

3. $\mathcal{P}$ as a convex subset of a Hilbert space.

The space $\mathcal{L}$ of random variables is a vector space. Define the norm of p by

(5) $$||p||^2 = \int p^2(x)d\mu(x).$$

Let $\mathcal{L}_2(\mu)$ be the space of all p where $||p|| < \infty$. Define an inner product on $\mathcal{L}_2(\mu)$ by

$$(6) \qquad < p, q > = \int p(x)q(x)d\mu(x).$$

The space $\mathcal{L}_2(\mu)$ together with the inner product given by (6) is a Hilbert space (Halmos (1957), §9). The following relation is crucial in applying convex analysis to conditions (1) and (2):

$$(7) \qquad < p, q > = E(q|p) \quad \text{whenever} \quad p \in \mathcal{P}.$$

To avoid difficulties in (1), (2), and (7) we will assume that the range of f is contained in the set

$$(8) \qquad \mathcal{L}_1 = \mathcal{L}_1(\mathcal{P}) = \{q \in \mathcal{L}\colon E(q|p) \text{ exists and is finite for all } p \in \mathcal{P}\}.$$

An important function related to the payoff function $f(p) \in \mathcal{L}_1$ is the expected payoff function defined on $\mathcal{P}$ by

$$(9) \qquad H(p) = < f(p), p > .$$

In terms of H, conditions (1) and (2) become, respectively,

$$(10) \qquad H(p) >(< p, f(q) >) \text{ if } p \neq q \text{ a.e. } \mu$$

and

$$(11) \qquad H(p) \geq(< p, f(q) >) \text{ for all } p, q \in \mathcal{P}.$$

Conditions (9), (10), and (11) can be expressed in terms of useful concepts found in the theory of convex analysis.

4. Review of some concepts of convex analysis.

Rockafellar (1970) gave definitions and theorems about convex functions and subgradients for the special case that the Hilbert space $\mathscr{K}$ is Euclidean. These concepts are sufficient for the finite discrete case of McCarthy's theorem. We have utilized some of Rockafellar's definitions and theorems for more general spaces when they apply. Rather than prove that convex functions are continuous on the interior of their domains we must assume it. Theorem 5 implies that the graph of any continuous convex function H over a nonempty open domain corresponds to the boundary of a convex set whose supporting hyperplanes correspond to subgradients of H. When $\mathscr{K}$ is Euclidean, Theorem 5 is implied by Theorem 23.4 of Rockafellar. Theorem 4 is given because, unlike Rockafellar, we have not assumed convexity as part of the definition of subgradient. The theory of convex sets in more general spaces is taken from Valentine (1964), Halmos (1957), and Köthe (1969).

Let $\mathscr{K}$ denote any Hilbert space. Then the space $\mathscr{K} \times R$ is a Hilbert space with inner product given by

$$(12) \qquad < (p, \alpha), (q, \beta)> = < p, q > + \alpha\beta.$$

The <u>epigraph</u> of a real-valued function H on a convex subset C of $\mathscr{K}$ will be denoted by epi (H) and defined by

$$\text{epi } (H) = \{(p, \alpha): p \in C, \alpha \in R, \alpha \geq H(p)\}.$$

H is a convex function iff epi (H) is a convex set. For this reason, the theory of convex sets, in Valentine (1964) for example, can be applied to convex functions. It is clear that the graph of H is contained in bdry(epi (H)) where the topology is given by the norm

(13) $\|(p, \alpha)\|^2 = \|p\|^2 + \alpha^2.$

Generalized concepts of tangency are those of supporting hyperplanes to a set in a topological linear space $\mathcal{L}$ and subgradients of functions on a Hilbert space $\mathcal{H}$. A hyperplane is a set $\{p \in \mathcal{L}: h(p) = \alpha\}$ where h is a linear functional on $\mathcal{L}$, $h \not\equiv 0$. The hyperplane $\{p \in \mathcal{L}: h(p) = \alpha\}$ supports a set C at $x_0 \in C$ if $h(x) \geq \alpha$ for all $x \in C$ and $h(x_0) = \alpha$. Theorem 1 below is implied by Theorems 2.15 and 4.1 of Valentine (1964). Valentine's Theorems 2.8 and 2.15 are incorrect since he does not state the interior of the convex set must be nonempty. Valentine's proofs however are valid since they depend on the correctly stated Theorem 2.7. An example of a convex set in which every point is a boundary point, yet no point except 0 has a hyperplane of support through it, is the space $\mathcal{L}_2^+(\mu)$ of Example 7, Section 5.

<u>Theorem 1</u>.

If C is a convex subset of a topological linear space $\mathcal{L}$ and if the interior of C is nonempty, then through each of its boundary points there passes a closed hyperplane of support. Conversely, if C is closed, if the interior of C is nonempty, and if through each of its boundary points there passes a plane of support, then C is convex.

If $\mathcal{L} = \mathcal{H}$ is a Hilbert space, then the hyperplanes of support of Theorem 1 can be characterized by using the Riesz representation theorem. We show this in Theorem 2.

Theorem 2.

If $\{p \in \mathscr{H}: h(p) = \alpha\}$ is a closed hyperplane contained in a Hilbert space $\mathscr{H}$, then there exists $q^* \in \mathscr{H}$ such that $h(p) = < p, q^* >$ for all $p \in \mathscr{H}$.

Proof:

The set $\{p \in \mathscr{H}: h(p) = \alpha\}$ is closed iff the linear functional h is continuous on $\mathscr{H}$ with $h \not\equiv 0$ (Theorem 2.12, Valentine (1964)). Continuity of h on the compact set $\{p: \|p\| \leq 1, p \in H\}$ implies there exists $M > 0$ such that $h(\frac{p}{\|p\|}) \leq M$. Hence by the Riesz representation theorem (page 31, Halmos (1957)), there exists $q^* \in \mathscr{H}$ such that $h(p) = < p, q^* >$ for all $p \in \mathscr{H}$. □

The supporting hyperplanes in $\mathscr{H} \times R$, with inner product given by (12), are of the form $\{(p, \alpha): h_1(p, 0) + h_1(0, \alpha) = \beta\}$ where h_1 is a linear functional on $\mathscr{H} \times R$. These hyperplanes are of two forms:

(i) $\{(p, \alpha): \alpha = \beta - h(p)\}$

(ii) $\{(p, \alpha): h(p) = \beta, \alpha \in R\}$.

The first set is a nonvertical hyperplane and the second is vertical. If H is convex on C then the supporting hyperplane at $(q, H(q))$ of epi (H) is seen to satisfy either

$$H(p) \geq h(p - q) + H(q) \quad p \in C \tag{14}$$

or

$$h(p) \geq \beta \quad p \in C, h(q) = \beta \tag{15}$$

where h is a linear functional on $\mathscr{H}$. (15) implies the only vertical supporting hyperplanes are on the boundary of C. The closed supporting hyperplanes in (i) which are given by linear functionals of the form $h(p) = < p, q^* >$ are of interest in this paper. The point q^* is a generalization of the gradient of H in R^n. We now generalize Rockafellar's definition of subgradient to the infinite-dimensional case.

Definition 1.

If H is defined on a convex set C contained in a vector space $\mathcal{L}$ and if there exists $q \in C$ and $q^* \in \mathcal{L}$ such that the inner product $< p, q^* >$ is defined for all $p \in C$ and

$$(16) \qquad H(p) \geq (< p - q, q^* >) + H(q) \quad \text{for all } p \in C$$

then q^* is a subgradient of H at q.

The following theorem shows that the subgradient is a generalization of the gradient when H is convex.

Theorem 3. (Theorem 25.1, Rockafellar (1970)).

Let H be a convex function on a convex set $C \subset R^n$. If H is differentiable at q, then $\nabla H(q)$ is the unique subgradient of H at q, so in particular

$$H(p) \geq (< p - q, \nabla H(q) >) + H(q) \quad \text{for all } p \in C.$$

Conversely, if H has a unique subgradient at q, then H is differentiable at q.

The following shows that the class of convex functions contains the functions which are "subdifferentiable" at each point in their domain.

Theorem 4.

If H has a subgradient q^* at each point q in a convex set C, then H is convex on C.

Proof:

Let $p \in C$, $q \in C$ and define $p_1 = (1-\lambda)p + \lambda q$. Let p_1^* be the subgradient of H at p_1. Then

$$H(p) \geq (< p - p_1, p_1^* >) + H(p_1)$$

and

$$H(q) \geq (< q - p_1, p_1^* >) + H(p_1).$$

Hence,

$$(1-\lambda)H(p) + \lambda H(q) \geq (< p_1 - p_1, p_1^* >) + H(p_1) = H(p_1). \square$$

The converse of Theorem 4 is not true in general. However, the converse is true whenever epi(H) contains an open set in $\mathcal{V} \times R$. This is given in Theorem 5. Lemma 1 restates the condition on epi(H) in terms of continuity. The proof follows directly from the definition of continuity.

Lemma 1.

Let C be a convex set in $\mathcal{V}$ whose interior is nonempty. Let H be a convex function on C which is continuous at a point $p \in \text{int}(C)$. Then epi (H) has a nonempty interior.

Theorem 5.

If C and H satisfy Lemma 1 then H has a subgradient $q^* \in \mathcal{V}$ at each point $q \in \text{int}(C)$.

Proof:

Theorems 1 and 2 imply that at every point $q \in C$ there exists a $q^* \in \mathcal{V}$ such that either (14) or (15) hold, where

$h(p) = < p, q^* >$. If $q \in \text{int}(C)$, then (15) cannot hold, so (14) holds and q^* is a subgradient of C. □

In Section 5 we give an example of a convex function H which has no subgradient at any point (Example 5), and an example of a continuous convex function whose subgradients are not members of $\mathscr{V}$ and $\text{int}(C) = \emptyset$ (Example 7). We also give an example where H does have a subgradient in $\mathscr{V}$ at each point in C, but H is not continuous (Example 6). In all the examples given, H is homogeneous.

If H is convex and homogeneous on a convex set C, then by putting $H(\lambda p) = \lambda H(p)$ the domain of H can always be extended to the <u>convex cone</u>

$$D = \{\lambda p: p \in C, \lambda > 0\}.$$

Thus without loss of generality we can assume H is defined on a convex cone.

<u>Lemma 2</u>.

If H is homogeneous on a convex cone D and if for each $q \in D$ there exists a subgradient $q^* \in \mathscr{V}$, then $H(q) = < q, q^* >$ for all $q \in D$.

<u>Proof</u>:

By letting $p_1 = \lambda p$, condition (16) and homogeneity imply

$$(17) \qquad H(p_1) \geq (< p_1 - \lambda q, q^* >) + \lambda H(q) \quad \text{for all } q, p_1 \in D.$$

Also condition (16) implies

$$(18) \qquad H(\lambda p) \geq (< \lambda p - q, q^* >) + H(q) \quad \text{for all } p, q \in D.$$

Taking the limit as $\lambda \to 0$ in (17) and (18) and letting $p_1 = q$ we obtain $H(q) = < q, q^* >$ for all $q \in \mathscr{V}$. □

Definition 2.

H is said to be strictly convex on a convex set C if

$$H((1-\lambda)p + \lambda q) < (1-\lambda)H(p) + \lambda H(q) \tag{19}$$

whenever $p \neq q$, $p \in C$, $q \in C$ and $0 < \lambda < 1$.

The following theorem is similar to page 94 of Valentine (1964).

Theorem 6.

The following statements are equivalent if H has a subgradient at each point in a convex set C:

(i) H is strictly convex;

(ii) H is nonlinear between any two distinct points in C;

(iii) Each nonvertical supporting hyperplane intersects epi (H) at exactly one point;

(iv) If $p \in C$ and $q \in C$ and $p \neq q$ then no subgradient of H at p is equal to a subgradient of H at q.

Proof:

(i) $\to$ (ii) $\longleftrightarrow$ (iii) $\longleftrightarrow$ (iv) are obvious. (ii) $\to$ (i) can be easily proved from the definition of convexity. □

The theory of this section is enough to prove the more general version of McCarthy's theorem, which we present in the next section. Some other concepts in convex analysis which can be applied to payoff functions will be given in Chapter III.

5. McCarthy's theorem.

McCarthy stated his theorem for the special case that $\mathscr{V} = R^n$ and $\mathscr{P}$ is the set of all n-dimensional probability vectors. The following is a restatement of McCarthy's theorem using our notation.

Theorem.

A payoff rule f satisfies (1) for all $p, p_1 \in \mathcal{P}$ iff there exists a homogeneous, convex function H such that $f(p)$ is the gradient of H at p. The function H satisfies (9).

The theorem is incorrect in that either condition (1) should be replaced by (2) or the convexity of H should be replaced by strict convexity. Also, f need not be continuous and H does not necessarily have a gradient at each point $p \in \mathcal{P}$. However, McCarthy does state that "the derivative has to be taken in a suitably generalized sense." Examples 3 and 4 are given below to illustrate McCarthy's conditions.

Example 3.

Let $f(p) = (c_1, c_2, \ldots, c_n)$. Then $H(p) = \sum_{k=1}^{n} c_k p_k$ satisfies McCarthy's theorem. Condition (2) holds but condition (1) doesn't because the payoff is independent of the forecaster's estimate.

Example 4.

Let $\mathfrak{X} = \{E, E^c\}$ and $p = (p_1, p_2)$ where $p_2 = 1 - p_1$.

Define

$$f(p) = (f_1(p), f_2(p)) = \begin{cases} (1, 0) & \text{if } p_1 > \frac{1}{2} \\ (0, 1) & \text{if } p_1 \leq \frac{1}{2} \end{cases}.$$

Then $H(p) = \langle p, f(p) \rangle = \max\{p_1, p_2\}$ is convex and homogeneous. However, H is not differentiable at $p_1 = p_2$. There is some leeway in defining $f_k(p)$ at $p_1 = p_2 = \frac{1}{2}$: $f_k(p)$ need only satisfy $0 \leq f_1(p) \leq 1$, $f_1(p) + f_2(p) = 1$ at $p_1 = \frac{1}{2}$.

Examples of strictly convex functions which are not differentiable everywhere can be found by considering functions of the form $H(p) = \max \{H_1(p), H_2(p)\}$ where the $H_k(p)$ are strictly convex on $\mathcal{P}$.

A corrected and generalized version of McCarthy's theorem is given below. Again, we assume $\mathcal{L}_1$ and $\mathcal{P}$ are subsets of $\mathcal{L}$ in which the inner product (expectation) $< p, q^* >$ is defined if $p \in \mathcal{P}$ and $q^* \in \mathcal{L}_1$.

Theorem 7.

A payoff rule f mapping $\mathcal{P}$ into $\mathcal{L}_1$ satisfies condition (2) [or condition (1)] iff there exists a homogeneous and convex [or strictly convex on $\mathcal{P}$] function H defined on the convex cone $D = \{\lambda p: p \in \mathcal{P}, \lambda > 0\}$ such that $f(p)$ is the subgradient of H at p for all $p \in \mathcal{P}$. The function H satisfies condition (9).

Proof:

Assume f satisfies (2), and define $H(\lambda p) = < \lambda p, f(p) >$ for $p \in \mathcal{P}, \lambda > 0$. Then $f(q)$ is a subgradient of H at q (condition (11)). If f satisfies (1) then no subgradient of H at p is a subgradient of H at q (condition (10)). Apply Theorem 4 for the convexity of H and Theorem 6 for strict convexity.

Conversely, if H has $f(q)$ as a subgradient at q, for each $q \in \mathcal{P}$, and if H is homogeneous and convex, then applying Lemma 2 and Definition 1, we obtain condition (2). Condition (iii) of Theorem 6 implies (1). □

When condition (1) does hold, H is strictly convex not on D but on $\mathcal{P}$. In general, the notions of strict convexity and homogeneity are contradictory.

It might be asked what class of functions H satisfy the conditions of Theorem 7. Although every function which is convex on $\mathcal{P}$ can be extended to a homogeneous and convex function on the cone D of Theorem 7, every such function does not satisfy the additional requirement of having subgradients at each point $q \in \mathcal{P}$. When $\mathcal{L} = R^n$, this additional requirement is met on the interior of $\mathcal{P}$. We will prove a more general result for $\mathcal{P} \subset \mathcal{L}_2(\mu)$ (Theorem 8). The following example shows that H must be restricted.

Example 5.

Let $\mathcal{P}$ be the class of continuous, bounded densities ($\sup_x p(x) < \infty$) on $(R, \mathcal{B}, \mu)$ where μ is Lebesgue measure and $\mathcal{B}$ consists of the Borel sets. Define $H(p) = \sup_x p(x)$. Then H is clearly convex on $\mathcal{P}$. However, H is neither continuous at any $p \in \mathcal{P}$ (with respect to $\|p\|$) nor does H have a subgradient for any $p \in \mathcal{P}$.

For the remainder of this section, the Hilbert space $\mathcal{V}$ can be taken to be either $\mathcal{L}_2(\mu)$ or the smallest closed subspace of $\mathcal{L}_2(\mu)$ containing $\mathcal{P}$, where $\mathcal{P} \subset \mathcal{L}_2(\mu)$. According to Theorem 8 below, the functions satisfying Theorem 8 are contained in the class of functions which are maximum expectations of payoff functions which encourage honesty.

Theorem 8.

If the set of densities $D \subset \mathcal{N}$ is a convex set and if H is convex and homogeneous on D and continuous at a point p in the interior of D, then there exists f such that conditions (9) and (11) hold on the interior of D. The range of f may be taken in $\mathcal{N}$.

Proof:

Whenever $p \in \mathrm{int}(D)$, apply Theorem 5 and let $f(p)$ be a subgradient of H at p. The proof follows from Lemma 2. □

The following theorem gives equivalent conditions on f for the conditions on H in Theorem 8.

Theorem 9.

If D is a convex cone in $\mathcal{N}$ whose interior is nonempty and if H and f satisfy (9) and (11) on D, then H is continuous at $p \in \mathrm{int}(D)$ iff there exists a neighborhood of p on which $\|f(\cdot)\|$ is bounded.

Proof:

Let $p_n \in D$, $\|p_n - p\| \to 0$ as $n \to \infty$. Let $q_n = \dfrac{f(p_n)}{\|f(p_n)\|^2}$. Then

$$H(p_n + q_n) \geq (\langle p_n + q_n, f(p_n) \rangle)$$

$$= H(p_n) + \langle q_n, f(p_n) \rangle$$

$$= H(p_n) + 1.$$

Thus, if H is continuous at p, we cannot have $\|q_n\| \to 0$. Thus we can not have $\|f(p_n)\| \to \infty$. Hence $\|f(\cdot)\|$ is bounded on a neighborhood of p.

Conversely, if $\|f(\cdot)\|$ is bounded on a neighborhood of p, then by the Cauchy-Schwarz inequality

(20) $\quad < p_n - p, f(p_n) > \to 0$ if $\|p_n - p\| \to 0$.

(9) and (11) imply

(21) $\quad H(p_n) \geq (< p_n, f(p) >)$

and

(22) $\quad H(p) \geq (< p, f(p_n) >)$.

(9) and (21) imply $\underline{\lim}\, H(p_n) \geq \lim < p_n, f(p) > = H(p)$. The latter equality follows from continuity of $< \cdot, f(p) >$ (Halmos,1957 §17).
(9), (20), and (22) imply

$$H(p) \geq \overline{\lim}\, H(p_n).$$

Hence $H(p) = \lim_{n \to \infty} H(p_n)$ if $\|p_n - p\| \to 0$. □

The results of Theorem 9 on local conditions on H and f easily give the following results on global conditions on H and f.

Corollary .

If f and H satisfy (9) and (11) on an open convex cone $D \subset \mathscr{X}$ then H is continuous on D iff $\|f(\cdot)\|$ is bounded on every closed set contained in D.

Proof:

Since $f(\lambda p) = f(p)$ if $\lambda > 0$, $\|f(\cdot)\|$ is bounded on every closed set contained in D is equivalent to $\|f(\cdot)\|$ bounded on every compact set in D which in turn is equivalent to the requirement of Theorem 9 that $\|f(\cdot)\|$ be "locally bounded" at each point $p \in \mathrm{int}(D)$. □

Theorem 10 shows that the restriction of the maximum expected payoff functions to be continuous with respect to $\|p\|$ on all of $\mathcal{L}_2(\mu)$ is merely the restriction of payoff functions to be bounded when the condition of encouraging honesty is met.

<u>Theorem 10</u>.

If H and f satisfy conditions (9) and (11) for all points in a Hilbert space $\mathcal{H}$, then the following are equivalent:

(i) H is continuous;

(ii) H is bounded on the sphere $\{p \in \mathcal{H}: \|p\| = 1\}$;

(iii) f is bounded.

<u>Proof</u>:

(ii) follows from (i) since the sphere is compact. Suppose condition (ii) holds. As a consequence of conditions (9) and (11) and the fact that $f(q) \in \mathcal{H}$ if $q \in \mathcal{H}$ we have

$$H\left(\frac{f(q)}{\|f(q)\|} \right) \geq \left\langle \frac{f(q)}{\|f(q)\|} , f(q) \right\rangle = \|f(q)\|$$

and hence (iii) follows. Condition (iii) implies (i) by the previous corollary. □

The following example shows that the class of maximum expected payoffs H satisfying the conditions of Theorem 6 contains discontinuous convex functions.

<u>Example 6</u>.

Let $\mathfrak{X} = [0, 1]$, μ be Lebesgue measure, and $\mathfrak{B}$ be the Borel subsets of $\mathfrak{X}$. Let $H(p)$ be the norm

$$H(p) = \left[\int |p(x)|^{\alpha} d\mu(x)\right]^{1/\alpha}$$

and define the payoff f to be

$$f(p) = [\frac{p(x)}{H(p)}]^{\alpha-1}$$

where $\alpha > 2$. Then H satisfies (9), and Holder's inequality is equivalent to (10), where $\mathcal{P}$ is taken to be the set of densities where H is finite. The domain of H is the familiar vector space $\mathcal{L}_\alpha(\mu) = \{p: H(p) < \infty\}$, and H is the usual norm. The space $\mathcal{L}_\alpha(\mu) \subset \mathcal{L}_2(\mu)$ is not closed with respect to the norm $\|p\|$, nor is the norm $H(p)$ continuous with respect to $\|p\|$.

Example 7, below, illustrates that Theorem 8 possibly can be generalized. The function H below cannot be extended to a convex domain whose interior is nonempty without losing its property of convexity; yet H satisfies (9) and (11) and is continuous.

Example 7.

Let $(\mathfrak{X}, \mathcal{B}, \mu)$ be the space given in Example 6. Let $\mathcal{L}_2^+ = \mathcal{L}_2^+(\mu) = \{p: p(x) \geq 0 \text{ for all } x \in \mathfrak{X}, \int p^2 d\mu < \infty\}$. Define f by

$$f(p)(x) = \begin{cases} \ln\left(\frac{p(x)}{\int p d\mu}\right) & \text{if } p(x) \neq 0,\ p \neq 0 \text{ a.e. } \mu \\ 0 & \text{if } p(x) = 0 \text{ or } p = 0 \text{ a.e. } \mu . \end{cases}$$

Define the function H by (9). Since

$$-e^{-1} \leq x \ln x \leq x^2$$

we have $pf(p)$ dominated above by $p^2(x)$ and below by $-e^{-1}$ if $\int p d\mu = 1$. Therefore, since μ is finite on $\mathfrak{X} = [0, 1]$, $H(p)$ is finite for all $p \in \mathcal{L}_2^+$. H and f do satisfy (9) and (11) on $\mathcal{L}_2^+$.

The function H is continuous on $\mathcal{L}_2^{+}$ with respect to $\|\cdot\|$. To prove this, let $p_n \in \mathcal{L}_2^{+}$, $p \in \mathcal{L}_2^{+}$ and $\|p_n - p\| \to 0$ as $n \to \infty$. Give $\varepsilon > 0$ and let $\delta > 0$ be such that $|x \ln x| < \varepsilon$ if $0 < x < \delta$ and such that $M \geq 1$ where $M = \frac{1}{\delta} \ln \frac{1}{\delta}$. Then it can easily be shown that

$$|x \ln x - y \ln y| \leq M|x^2 - y^2|$$

if either $x \geq \delta$ or $y \geq \delta$. Thus, by integrating $|p \ln p - p_n \ln p_n|$ over the set $\{x: 0 \leq x \leq 1,\ p(x) < \delta,\ p_n(x) < \delta\}$ and the set $\{x: 0 \leq x \leq 1,\ p(x) \geq \delta \text{ or } p_n(x) \geq \delta\}$ we obtain

$$|H(p) - H(p_n)| \leq 2\varepsilon + M \int_0^1 |p^2(x) - p_n^{\,2}(x)|dx.$$

Since $\lim_{n\to\infty} \int_0^1 |p^2(x) - p_n^{\,2}(x)|dx = 0$ and since $\varepsilon > 0$ is arbitrary, $\lim_{n\to\infty} H(p_n) = H(p)$.

If $p \in \mathcal{L}_2^{+}$, then every neighborhood of p contains functions which are negative over sets of positive measure. Hence $\mathcal{L}_2^{+}$ has no interior points and Theorem 7 does not apply.

Although every member of $\mathcal{L}_2^{+}$ is a boundary point, the space $\mathcal{L}_2^{+}$ has no closed hyperplanes of support through any of its nonzero members. For, if $(<\lambda q, q^* >) \geq (< q, q^* >)$ for all $\lambda > 0$ then $< q, q^* > = 0$, and if $(< p, q^* >) \geq 0$ for all $p \in \mathcal{L}_2^{+}$ then $q^* = 0$ a.e. μ. Thus the convex set $\mathcal{L}_2^{+}$ is a counterexample to the set C in Theorem 2.15 of Valentine (1964), unless we add the requirement that the interior of C be nonempty.

CHAPTER III

Support Functions

1. Introduction.

The viewpoint of a payoff function f to a forecaster, as in Chapter I, can be studied from the viewpoint of a class C of choices of payoffs to a forecaster, as in Chapter IV. The present chapter studies the one-to-one relationship between closed convex classes C and the corresponding expectation function H.

Let f and H satisfy conditions (9) and (11) on a subset $\mathcal{P} \subset \mathcal{H}$ of a Hilbert space. Let C_1 be the range of f. If when offered the payoff $f(p)$ for his appraisal p, a forecaster is allowed to be dishonest, then the forecaster is actually offered a choice between the random variables in C_1. Define the closed and convex set C by

$$(23) \qquad C = \{q \in \mathcal{H}:(< q, p >) \leq \sup_{q^* \in C_1} < q^*, p > \text{ for all } p \in \mathcal{P}\}.$$

C is closed because it is the intersection of the closed half-spaces $\{q \in \mathcal{H}: (< q, p >) \leq \sup_{q^* \in C_1} < q^*, p >\}$. Obviously C_1 is contained in C. In a sense, C and C_1 are equivalent because we could offer a forecaster his choice of the random variables in C and his maximum expectation would remain unchanged. For this reason, in this chapter we will take any class of choices of payoffs to be closed and convex.

Conversely, it will be shown in Section 2 that for any given closed convex C there corresponds an f and an H satisfying

(9) and (11), provided $\sup_{q \in C} < p, q >$ is attained in C for each $p \in \mathcal{P}$.

After the forecaster has made his choice $p_1^* \in C$, the class of possible values of the forecaster's undisclosed p is the set

(24) $$A^*(p_1^*) = \{p \in \mathcal{P}: < p_1^*, p > = \sup_{q \in C} < q, p > \}.$$

If p_1^* is the honesty encouraging payoff $f(p_1)$ then $A^*(p_1^*)$ is the class $A(p_1)$ given by (4). The sets $A^*(p_1^*)$ will be shown to be convex (the intersection of $\mathcal{P}$ with the normal cone of C at p_1^*) if C is convex and closed.

Dual relationships between H and C are given in this chapter. It is shown that p must be normal to C at $f(p)$ if $f(p) \in C$ is to encourage honesty. The theorems and definitions are found in Valentine (1964) or Rockafellar (1970), at least when $\mathcal{H}$ is Euclidean. Theorem 12 is given in Köthe (1969).

2. Support functions.

The maximum expectation function H of a payoff function f which encourages honesty, which was shown in section 5 to be convex and homogeneous, can be interpreted as a support function of a certain convex set. The theory of support functions given in this section is found in Part V of Valentine (1964). We again assume $\mathcal{H}$ is a Hilbert space, usually $\mathcal{L}_2(\mu)$.

Definition 3.

If C is a convex set then the function H defined on the set

$$(25) \qquad D = \{p \in \mathcal{H}: \sup_{q \in C} (< p, q >) < \infty\}$$

by the equation

$$(26) \qquad H(p) = \sup_{q \in C} < p, q >$$

is the <u>support function</u> of C.

The following theorem is a direct result of (25) and is Theorem 5.1 of Valentine (1964).

<u>Theorem 11</u>.

The domain D of definition of a support function H is a convex cone.

Note that (26) implies H is homogeneous and convex on D. We defined H in terms of C in (26). There is a one-to-one correspondence between homogeneous and convex functions H and closed convex sets C, where the domain of H is given by (25). We will prove this in Theorem 13.

For a given homogeneous and convex function H defined on a set $D \subset \mathcal{H}$, define C by the equation

$$(27) \qquad C = \{p^*: (< p, p^* >) \leq H(p) \quad \text{for all} \quad p \in D\}.$$

Then C is closed because the half-space $\{p^*: (< p, p^* >) \leq H(p)\}$ is closed for each $p \in D$. We will prove Theorem 13 from the following theorem given in Köthe (1969), §20.7, Theorem 1.

<u>Theorem 12</u>.

Let C be a closed convex subset of a locally convex space L and let K be a compact convex set which is disjoint from A.

Then there exists a closed hyperplane which separates A and K strictly.

Theorem 13 specializes to Theorem 5.3 of Valentine (1964) whenever $\mathcal{H}$ is Euclidean.

Theorem 13.

Let C be a nonempty closed convex set contained in a Hilbert space $\mathcal{H}$ and let H be its support function with nonempty domain of definition D. Then C satisfies (27).

Proof:

Let C_1 be given by (27) where H is the support function of C. Then $C \subset C_1$. Suppose there exists $q \in C_1 - C$. Since $\{q\}$ is compact, Theorem 12 and Theorem 2 imply there exists $p \in \mathcal{H}$ and $\alpha \in R$ such that $(< p, q >) > \alpha$ and $(< p, p^* >) < \alpha$ for all $p^* \in C$. Equation (25) implies $p \in D$ and (26) implies $H(p) \leq \alpha < (< p, q >)$. This contradicts $q \in C_1$. Thus $C_1 = C$. □

A theorem similar to that given by Theorem 7 can be given in terms of the closed convex set C in (27). To state the theorem we need the following two definitions.

Definition 4 (pages 15 and 100, Rockafellar (1970)).

Let C be a closed convex set. Then p is normal to C at p^* if

$$< p, p^* > = \sup_{q \in C} < p, q >.$$

Definition 5 (Definition 7.5, Valentine (1964)).

A convex set $C \subset \mathcal{H}$ is smooth if at each of its boundary points there is a unique hyperplane of support to C. A convex function H is <u>smooth</u> (differentiable if $\mathcal{H} = R^n$) if H has a unique subgradient at every point of its domain.

Theorem 14.

A payoff function f on a space $\mathcal{P} \subset \mathcal{H}$ encourages honesty [strictly] if and only if there exists a closed convex [smooth] set $C \subset \mathcal{H}$ such that for each $p \in \mathcal{P}$, p is normal to C at f(p). The set C may be taken to be

$$C = \{q: (< p, q >) \leq (< p, f(p) >) \text{ for all } p \in \mathcal{P}\}. \tag{28}$$

Proof:

If f encourages honesty, define C by (28). By definition, p is normal to C at f(p). Conversely, if C is closed and convex and p is normal to C at f(p), then $f(p) \in C$ (else C is not closed) so

$$(< p, f(p) >) \geq (< p, f(q) >) \text{ for all } p, q \in \mathcal{P}. \tag{29}$$

Equality holds in (29) for some p and $q \in \mathcal{P}$, where $p \neq q$, iff both p and q are normal to C at f(q). In other words, C has two hyperplanes of support through f(q). Thus if (28) and (29) hold then C is smooth iff strict inequality holds in (29) for all $p, q \in \mathcal{P}$, $p \neq q$. □

As in the proofs of Theorem 7 and Theorem 14, a closed convex set C is smooth iff its support function H is strictly convex. This relationship is dual: C contains a line segment on its boundary (is not strictly convex) iff there exists p which is normal to C at two distinct points f(p) and p^*, and this is true iff H is not smooth (there exists p such that H has two subgradients $f(p)$ and p^* at p).

The question, about which Theorem 8 is concerned, of when a convex homogeneous function H has a subgradient at each point $p \in \mathcal{P}$, can be restated as the question of when a closed convex set C has a closed hyperplane of support of the form $\{p^*: < p, p^* > = \alpha\}$ for each $p \in \mathcal{P}$. The following theorem, which is similar to Theorems 7 and 8, is given in Valentine (1964), Theorem 5.2.

Theorem.

If C is a bounded closed nonempty convex set in $\mathcal{H}$, then for each $p \in \mathcal{H}$ there exists a point $p^* \in C$ such that $H(p) = < p, p^* >$ where H is the support function of C.

The following example shows that the boundedness of C in the above Theorem is necessary, even when its support function is bounded.

Example 8.

Let $\mathcal{H} = R^2$. Let $C = \{(p_1, p_2): p_1 < 0, p_2 \leq \frac{1}{p_1}\}$. Then the boundary of C is the graph of the function $h(x) = \frac{1}{x}$ for $x < 0$. C is closed and convex. The set D given by (25) is the set $\{p: p_1 \geq 0, p_2 \geq 0\}$. If $p \in D$ and $p_1 \neq 0, p_2 \neq 0$ then p is normal to C at the point $f(p) = (-\sqrt{\frac{p_2}{p_1}}, -\sqrt{\frac{p_1}{p_2}})$. The support function H is defined for all $p \in D$ by $H(p) = -2\sqrt{p_1 p_2}$. H is bounded, continuous and convex on the set $\mathcal{P} = \{p \in D: p_1 + p_2 = 1\}$, yet H has no subgradients at the points $(0, 1)$ or $(1, 0)$.

The set $A(p_1)$ given in equation (4) of Chapter 2 is the intersection of $\mathcal{P}$ with a normal cone, defined below.

Definition 6 (page 135, Valentine (1964)).

Let C be a closed, convex set. Let $Q(p^*)$ be the set of all points which are normal to C at p^*. Then $Q(p^*)$ is the normal cone of C at p^*.

If f satisfies (2) and C satisfies (27), then $A(p_1) = Q(f(p_1)) \cap \mathcal{P}$ for each $p \in \mathcal{P}$. According to Theorem 11.1 of Valentine (1964), the normal cone $Q(p^*)$ is convex for each p^* on the boundary of C. Thus $A(p_1)$ is convex for each $p_1 \in \mathcal{P}$. Alternatively, if (9) and (11) hold, then $A(p_1)$ is the projection on $\mathcal{H}$ of the intersection of the convex set epi $(H) \subset \mathcal{H} \times R$ with its supporting hyperplane $\{(p, \alpha): \alpha = < p, f(p_1) > \}$. Thus $A(p_1)$ is convex.

In the following example, f, H, and C satisfy conditions (9), (11), and (27). The convex cones $Q(p^*)$ are illustrated in Figure 1 when they consist of more than one point—when the boundary of C is not smooth at p^*. Figures 1 and 2 illustrate the relationships between C and f and between H and f.

Example 9.

Let $\mathcal{H} = R^2$. Let

$$C = \{p: 0 \leq p_1 \leq p_2, \|p\| \leq 1\} \cup \{p: 0 \leq p_2 \leq p_1, p_1 + p_2 \leq \sqrt{2}\}.$$

Then C is a closed convex set. If p is normal to C at $f(p)$, and if p is not of the form $(0, -a)$, $(-a, 0)$, or (a, a) where $a \geq 0$, then $f(p)$ is uniquely determined:

$$
f(p) = \begin{cases} \frac{p}{\|p\|} & \text{if } 0 \le p_1 < p_2 \\ (\sqrt{2}, 0) & \text{if } 0 < p_1, \; p_2 < p_1 \\ (0, 0) & \text{if } p_1 < 0, \; p_2 < 0 \\ (0, 1) & \text{if } p_1 \le 0, \; p_2 > 0. \end{cases}
$$

The support function H of C is defined by f in equation (9):

$$
H(p) = \begin{cases} \|p\| & \text{if } 0 \le p_1 \le p_2 \\ \sqrt{2} \; p_1 & \text{if } 0 \le p_1, \; p_2 \le p_1 \\ p_2 & \text{if } p_1 \le 0, \; p_2 \ge 0 \\ 0 & \text{if } p_1 \le 0, \; p_2 \le 0. \end{cases}
$$

Note the dual relationship between smoothness and strict convexity of H and C. Also note that f(p) lies on the boundary and is continuous where the boundary of C is strictly convex, and constant on the cone where the boundary is not smooth (rough). The boundary of C differs from the range of f only at the points where C is not strictly convex. See Figures 1 and 2.

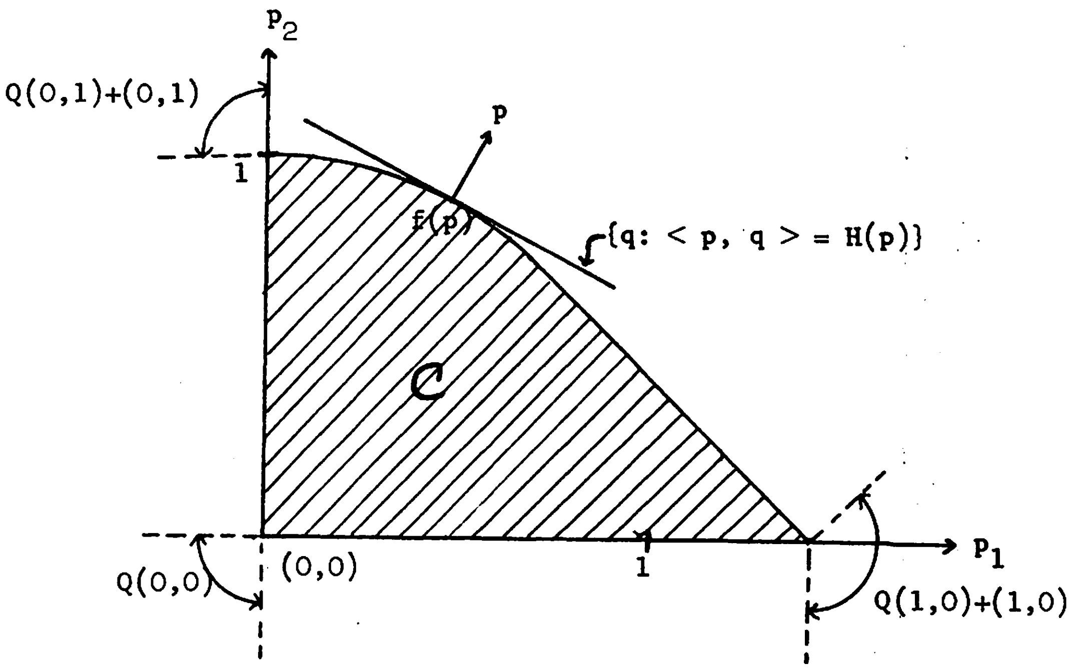

Figure 1. The convex set C of Example 9 and its cones of support.

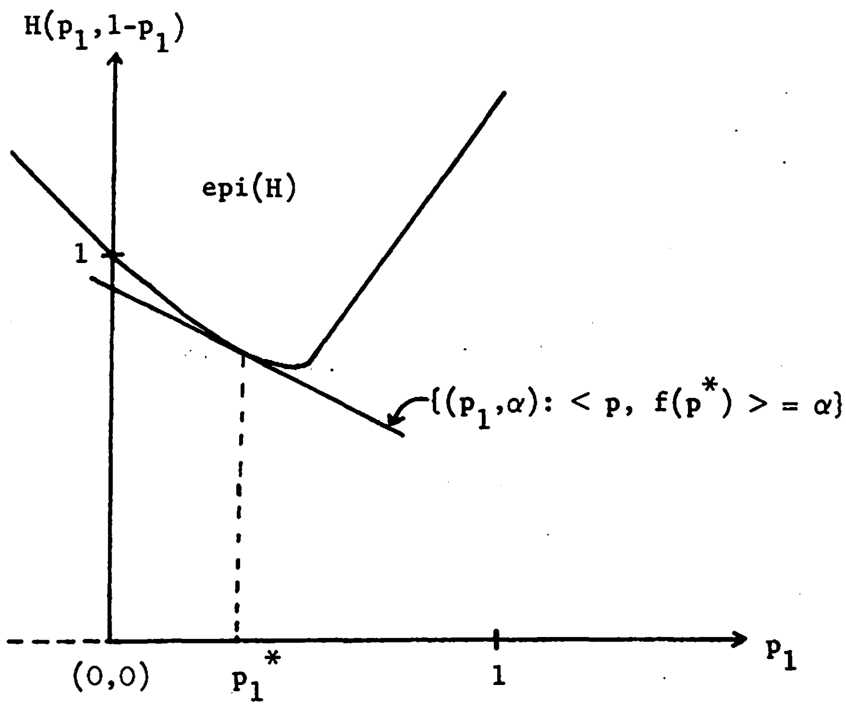

Figure 2. The support function H as a function of p_1, on the set $\{p: p_1 + p_2 = 1\}$.

CHAPTER IV

Some Sequential Procedures Which Encourage Honesty

1. Introduction.

Definitions of "subjective" or "personal" probability, such as that of Savage, involve hypothetical preferences or choices of a subject which carry information about subjective probability values. For example, Savage (1954, p. 28) says:

> We therefore address him thus: "We see you are about to open those eggs. If you will be so cooperative as to guess that one or the other egg is good, we will pay you a dollar, should your guess prove correct. If incorrect you and we are quits, except that we will in any event exchange your two eggs for two of guaranteed goodness." If under these circumstances the person stakes his chances on the brown egg, it seems to me to correspond well with ordinary usage to say that it is more probable to him that the brown one is good than the white one is.

By replacing the breaking of one egg (outcomes "good" or "bad") with the toss of a fair coin (outcomes "heads" or "tails") we may determine whether the subjective probability of the remaining egg being good is $\geq 1/2$ or $\leq 1/2$. If further choices are then offered, the subjective probability presumably can be restricted to narrower limits. We find however that certain difficulties arise in the choice of stake at each stage of the procedure. In short, an apparently "dishonest choice" may in some cases be advantageous to the subject. The present chapter is concerned with the characterization of procedures which "encourage honesty."

Sequential procedures are studied in which, for simplicity, each choice divides in half the range of possible values of the probability. These procedures are described in Section 2. Section 3 gives expressions for the cumulated expected payoff. In Section 4 we define payoffs which "encourage honesty" to be those having the property that the "honest choice" by the subject at each stage gives the largest cumulated expected payoff when the true probability is known. Sections 5 and 6 establish some properties of expected payoffs, including characteristics of payoffs which encourage honesty for finite sequential procedures. The infinite case is discussed in Section 7.

2. Description of the procedure.

Let A denote an event having unknown probability p and let B_r denote an event with known probability r. (We assume such exist for $r = 1/2, 1/4, 3/4, 1/8, 3/8, 5/8$, etc. Of course the existence of a single fair coin which can be tossed repeatedly implies this.) The appraiser is offered a sequence of choices of prospects. At any step Prospect A is a payoff of $g(r)$ if A occurs and Prospect B is a payoff of $g(r)$ if B_r occurs. It remains to describe the sequence $\{r_n\}$ of r-values and to choose $g(r)$. This is to be done in such a way that the best choices for the appraiser when p is actually known will allow us to deduce the value of p from his choices.

Consider the following particular choice. At Step 1, $r = r_1 = 1/2$. At Step 2, $r = r_2 = 1/4$ or $3/4$ according as Prospect B or A was chosen at Step 1. Similarly at Step n, $r = r_n$ will be either $r_{n-1} - 2^{-n}$ or $r_{n-1} + 2^{-n}$ according as the preceding choice was B or A. Clearly the idea here is to obtain a sequence of

r values converging to the true p when it is known, or to the appraised value $\hat{p}$ in any case. The appraiser will presumably attempt to maximize his expected payoff which he will calculate using $\hat{p}$ rather than p when the latter is unknown. Of course at Step 1 he then chooses A if $\hat{p} > 1/2$ provided he only considers the payoff determined by this first choice. If however the rules for the subsequent steps are known to him, the "dishonest" choice of B when $p > 1/2$ could actually increase his expectation at later stages. For this reason the choice of values $g(r)$ is relevant to an honest appraisal.

3. Expressions for expected payoff functions.

Let $a_j = -1$ if Prospect B is chosen at the jth step and $a_j = +1$ if A is chosen. Then r_n can be expressed by

$$r_n = \frac{1}{2} + \sum_{j=1}^{n-1} 2^{-j-1} a_j, \qquad n = 2, 3, \ldots, \tag{30}$$

and $r_1 = 1/2$.

There are slightly troublesome difficulties which arise when the appraiser is indifferent between the two prospects. For example, if $p = 1/2$ he is indifferent between A and B. However in the case when $n \to \infty$ we can still obtain $r_n \to 1/2$ whether $r_2 = 1/4$ or $3/4$. It would be desirable here to choose the values $g(r)$ so that the expected payoff does not depend on which choice is made in any case where the appraiser is indifferent. The extent to which this can be accomplished will be considered below.

It is convenient to define

(31) $$C_n = \{x \mid x = 2^{-n}k,\ k = 1, 3, \ldots, 2^n - 1\}$$

(32) $$D_n = \bigcup_{j=1}^{n} C_j$$

(33) $$D = D_\infty .$$

Then our payoff function $g(r)$ is to be defined for all $r \varepsilon D$ (although we will also consider finite procedures requiring only $r \varepsilon D_n$).

At Step n the payoff is either $g(r_n)$ with probability r_n if B is chosen $(a_n = -1)$ or $g(r_n)$ with probability p if A is chosen $(a_n = +1)$. Thus the increment in the expected payoff is $r_n g(r_n)$ if $a_n = -1$ or $pg(r_n)$ if $a_n = +1$. The cumulated expected payoff at the <u>nth</u> step depends on p and on the partial sequence of choices

(34) $$\mathcal{G}_n = (a_1, a_2, \ldots, a_n)$$

and can be expressed as

(35) $$H_n(\mathcal{G}_n, p) = F_n(\mathcal{G}_n) + pG_n(\mathcal{G}_n)$$

where

(36) $$F_n(\mathcal{G}_n) = \sum_{\substack{1 \le k \le n \\ a_k = -1}} r_k g(r_k)$$

(37) $$G_n(\mathcal{G}_n) = \sum_{\substack{1 \le k \le n \\ a_k = +1}} g(r_k).$$

Now let r be any value $0 < r < 1$ and $r \notin D_n$. Then r determines a unique binary expansion up to the nth term, and so we may replace (34) by

$$H_n(r, p) = F_n(r) + pG_n(r) \tag{38}$$

understanding that r determines a unique G_n. From (38) we see that the cumulated expected payoff $H_n(r, p)$ is linear in p and a step function in r whose discontinuities occur at the points of D_n, that is at multiples of 2^{-n}, or we may say at exactly those points of $(0, 1)$ where $H_n(\cdot, p)$ is not defined.

It is straightforward to calculate the jump in $H_n(\cdot, p)$ at each discontinuity. Consider first Case 1: $r = k/2^n$ where k is odd (that is $r \in C_n$). We find

$$F_n(r-) - F_n(r+) = rg(r) \tag{39}$$

$$G_n(r+) - G_n(r-) = g(r). \tag{40}$$

For Case 2, we express any other point r of discontinuity ($r \in D_{n-1} = D_n - C_n$) uniquely in the form $r = 2^{-s}k$ where k is odd and $s < n$. For these we find

$$F_n(r-) - F_n(r+) = rg(r) - \sum_{j=1}^{n-s} (r + 2^{-s-j})g(r + 2^{-s-j}) \tag{41}$$

$$G_n(r+) - G_n(r-) = g(r) - \sum_{j=1}^{n-s} g(r - 2^{-s-j}). \tag{42}$$

4. Payoffs which encourage honesty.

The procedure described thus far can be related to Chapters I and II. Define $h^{(n)} = (h_1^{(n)}, h_2^{(n)})$ by

$$(43) \qquad h_1^{(n)}(r) = \begin{cases} g(r_n) & \text{if } a_n = +1 \\ g(r_n)r_n & \text{if } a_n = -1 \end{cases}$$

$$(44) \qquad h_2^{(n)}(r) = \begin{cases} 0 & \text{if } a_n = +1 \\ g(r_n)r_n & \text{if } a_n = -1 \end{cases}$$

and define $f^{(n)} = (f_1^{(n)}, f_2^{(n)})$ by

$$(45) \qquad f^{(n)}(r) = \sum_{k=1}^{n} h^{(k)}(r).$$

Then at each n, $h^{(n)}$ is a payoff function which encourages honesty (satisfies condition (2)) because $r > r_n$ implies $a_n = +1$ and $r < r_n$ implies $a_n = -1$. We wish to consider procedures in which for any fixed n the cumulated payoff function $f^{(n)}$ encourages honesty.

The function $H_n(r, p)$ satisfies for each n,

$$(46) \qquad H_n(r, p) = < f^{(n)}(r), p > .$$

Thus for given n, $H_n(p, p)$ is the function $H_n(p)$ defined by (9). Hence if $f^{(n)}$ is to encourage honesty, it is necessary and sufficient that $H_n(p, p)$ be convex on $(0, 1)$ [or $\rho \subset [0, 1]$].

We assume that the appraiser knows the rules by which the payoffs are to be made, including the payoff values $g(r)$. The range of values of $H_n(r, p)$ (where $0 < p < 1$ is fixed and

$r \in (0, 1) - D_n)$ is the set of all possible expected payoffs for the appraiser at time n. From (35) through (38) the appraiser can calculate his expected payoff from the value of r which determines the sequence G_n and from a presumed value of p.

A payoff function would strictly encourage honesty if

$$(47) \qquad H_n(p, p) > H_n(r, p) \quad \text{for all} \quad r \neq p.$$

But since $H_n(r, p)$ is a step function in r we cannot actually achieve inequality for all $r \neq p$. If $p \neq 2^{-n}k$ (k odd), let I_{kn} be the interval $(2^{-n}k, 2^{-n}(k+1))$ (k odd) which contains p. We require

$$(48) \qquad H_n(p, p) > H_n(r, p) \quad \text{all} \quad r \notin I_{kn},$$

for all $p \neq 2^{-n}k$. Any payoff $g(r)$ such that (48) is satisfied will be said to "encourage honesty" at the <u>nth</u> stage. For p values on the boundary points $2^{-n}k$ no additional requirement is needed since (48) is strong enough to imply appropriate behavior. More specifically, it can be shown that when $p = 2^{-n}k$

$$(49) \qquad H_n(p+, p) = H_n(p-, p) \geq H_n(r, p) \quad \text{all} \quad r \in [0, 1] - D_n.$$

The proof uses (48) and continuity of $H_n(r, p)$ as a function of p. Hence if (48) holds and $p \in D_n$ then the appraiser's maximum expected payoff is $H_n(p+, p) = H_n(p-, p)$. For simplicity, extend the domain of definition of $H_n(r, p)$, $F_n(r)$, and $G_n(r)$ to $r \in (0, 1)$ by right-continuity. Then (48) is equivalent to

(50) $$H_n(p, p) \geq H_n(r, p)$$

with equality only when both p and r belong to the same interval $(2^{-n}k, 2^{-n}(k+1))$ or when $p = 2^{-n}k$ and $r \in (2^{-n}(k-1), 2^{-n}(k+1))$.

5. Properties of cumulated honesty encouraging payoffs.

We first consider properties of H_n defined by (38) which hold for rather arbitrary F_n and G_n. Let F and G be defined on $(0, 1)$ and define

(51) $$H(r, p) = F(r) + pG(r) \quad \text{for} \quad r, p \in (0, 1)$$

and let

(52) $$H(p, p) \geq H(r, p) \quad \text{for} \quad r, p \in (0, 1).$$

Except that the payoff function $(G(p) - F(p), F(p))$ has its range in R^2 and its domain in R^1, the discussion of this section is a special case of Chapter I, in particular Theorem 7.

Theorem 15 below is proved directly from (51) and (52). Condition (ii) points out why "encouraging honesty" provides an "incentive for accuracy."

Theorem 15.

Let F, G, and H be defined on $(0, 1)$ and satisfy (51) and (52). Then (i) F is nonincreasing and G is nondecreasing. (ii) for fixed p, $H(r, p)$ is nonincreasing for $r > p$ and nondecreasing for $r < p$. (iii) $H(p, p)$ is continuous.

Proof:

(i) Using (52) twice we have

$$\begin{aligned} F(p) + G(p)p &\geq F(r) + G(r)p \\ &= F(r) + G(r)r - G(r)(r-p) \\ &\geq F(p) + G(p)r - G(r)(r-p) \\ &= F(p) + G(p)p + G(p)(r-p) - G(r)(r-p). \end{aligned}$$

Thus $(G(p) - G(r))(r-p) \leq 0$, which implies G is nondecreasing. Assuming $p < r$, using (52) again, and $G(p) \leq G(r)$

$$\begin{aligned} F(p) + pG(p) &\geq F(r) + pG(r) \\ &\geq F(r) + pG(p) \end{aligned}$$

completing the proof of (i). (ii) If $p \geq r \geq s$, by (52) and (i),

$$\begin{aligned} F(r) + pG(r) &= F(r) + rG(r) + (p-r)G(r) \\ &\geq F(s) + rG(s) + (p-r)G(s) \\ &= F(s) + pG(s). \end{aligned}$$

If $p \leq s \leq r$

$$\begin{aligned} F(r) + pG(r) &= F(r) + sG(r) + (p-s)G(r) \\ &\leq F(s) + sG(s) + (p-s)G(s) \\ &= F(s) + pG(s) \end{aligned}$$

which proves (ii). (iii) Let $\{y_n\}$ be either an increasing or a decreasing sequence with $\lim y_n = p$. Then both $\lim F(y_n)$ and $\lim G(y_n)$ exist, and using (52) twice

$$\begin{aligned} F(p) + pH(p) \geq \lim \{F(y_n) + pG(y_n)\} &= \lim \{F(y_n) + y_nG(y_n)\} \\ &\geq \lim \{F(p) + y_nG(p)\} \\ &= F(p) + pG(p). \; \square \end{aligned}$$

Theorem 16.

Let $F(\cdot)$ and $G(\cdot)$ be defined on $(0, 1)$ and let H be defined by (51). Then (52) holds if and only if G is nondecreasing and

$$F(r) = \int_r^1 z\, dG(z) + c \tag{53}$$

for some constant c.

Proof:

Assume (53) and G nondecreasing. Then

$$H(r, r) - H(s, r) = \begin{cases} \int_r^s (z-r)dG(z) \geq 0 & \text{if } s > r \\ \int_s^r (r-z)dG(z) \geq 0 & \text{if } s < r \end{cases}$$

so that (52) holds. Now assume (52). Then by Definition 1, $G(r)$ is a subgradient of $H(p, p)$ at $p = r$ for every $r \in (0, 1)$. Hence H is convex by Theorem 1. Since convexity implies absolute continuity on $(0, 1)$, Theorem 3 implies

$$H(p, p) = \int_0^r G(z)dz + c. \tag{54}$$

By integration by parts, (51) implies (53) and (54) are equivalent expressions. □

Corollary 1.

If F and G are right-continuous step functions, then necessary and sufficient conditions for (52) are: (i) G is nondecreasing; (ii) $H(p, p)$ is continuous for all p.

Proof:

Let $\{r_k\}_{k=1}^{\infty}$ be the points where either F or G is discontinuous. Then a necessary and sufficient condition for (53) is

$$F(r_k-) - F(r_k+) = r_k[G(r_k+) - G(r_k-)] \quad \text{all } k,$$

and this is equivalent to (ii). □

The convex sets $A(p_1)$ defined by equations (4) and (24) are defined in the present context by

$$A(r) = \{p: H(p, p) = H(r, p)\}.$$

Corollary 2 below states that these sets are the closures of the intervals where F or G are constant. Define the interval $I_G(r)$ by

$$I_G(r) = \{p: G(r) = G(p-) \text{ or } G(r) = G(p+)\}.$$

Corollary 2.

If (51) and (52) hold then

$$A(r) = I_F(r) = I_G(r).$$

Proof:

If (51) and (52) hold then $p \in A(r)$ iff

$$F(p) - F(r) = p[G(r) - G(p)]. \tag{55}$$

Theorems 15 and 16 imply G is nondecreasing and (55) holds iff

$$\int_p^r z dG(z) = \int_p^r p dG(z), \text{ or}$$

$$\int_p^r (z-p) dG(z) = 0.$$

But the above integral is zero iff $G(z) = G(r)$ for all z between r and p. Hence $A(r) = I_G(r)$. Theorem 16 implies $I_F(r) = I_G(r)$. □

6. Properties of payoffs which encourage honesty.

We now apply the preceding results to determine properties of the payoffs $g(r)$ which will encourage honesty.

Lemma 3.

Let F_n, G_n, H_n be defined by (38) through (42). Then $H_n(p, p)$ is continuous if and only if

$$(56) \qquad \sum_{j=1}^{n-s} (r+2^{-s-j})g(r+2^{-s-j}) = r \sum_{j=1}^{n-s} g(r-2^{-s-j}) \quad \text{for all } r \in C_s$$

and for all $s = 1, 2, \ldots, n-1$.

Proof:

$H_n(p, p)$ is continuous at $r \in C_n$ by (39) and (40). At $r \in D_{n-1} = D_n - C_n$, continuity of $H_n(p, p)$ is equivalent to $F_n(r-)-F_n(r+) = r(G(r+)- G(r-))$. (41) and (42) give (56). □

Theorem 17.

Let F_n, G_n, H_n be defined by (38) through (42) and the assumption of right-continuity. Then necessary and sufficient conditions for (48) are (56) and

$$(57) \qquad g(r) > \sum_{j=1}^{n-s} g(r-2^{-s-j}) \quad \text{all } r \in C_s,$$

for $s = 1, 2, \ldots, n-1$.

Proof:

We can use the theorem of Section 5 with subscript n added to F, G, and H. Using Corollary 3 it is easy to see that (48) is equivalent to (52) with the additional condition that $G_n(p+) \neq G_n(p-)$ for $p \in D_n$. Hence by Corollaries 2 and 3 we have (48) holds if and only if the following three conditions hold:

(i) G_n is nondecreasing

(ii) $H_n(p, p)$ is continuous

(iii) $G_n(p+) \neq G_n(p-)$ for $p \in D_n$.

By Lemma 3, condition (ii) is equivalent to (56). Since G_n is a step function with jumps given by (42), we have conditions (i) and (iii) equivalent to (57). □

Theorem 18.

Let F_n, G_n, H_n be defined by (38) through (42) for $n = 1, 2, \ldots$. Then necessary and sufficient conditions for (48) to hold for all positive integers n are that for every $k = 3, 5, \ldots, 2^m-1$,

$$(58) \qquad g(2^{-m}k) = \frac{k-1}{k}\, g(2^{-m}(k-2)) = h(k)g(2^{-m}) \quad m = 2, 3, \ldots,$$

$$(59) \qquad g(2^{-m}) \geq \frac{1}{h(k)} \sum_{j=1}^{\infty} h(2^j k-1)g(2^{-m-j}) \quad m = 1, 2, \ldots,$$

where

$$(60) \qquad h(k) = \frac{1}{2}\,\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}(k+1)\right)/\Gamma\left(\frac{1}{2}k+1\right) = \frac{2\cdot 4\cdots(k-3)(k-1)}{1\cdot 3\cdots(k-2)k} .$$

Proof:

Suppose (58) and (59) hold. For given r, s, j, define k by $k = 2^m r + 1$ where $m = s + j$. Then $2^{-m}k = r + 2^{-s-j}$,

$2^{-m}(k-2) = r - 2^{-s-j}$, and (58) implies termwise equality of the sums in (56). To prove (57), for given $r \in C_s$ define m, k by $m = s$ and $2^{-m}k = r$. Then by (58) $h(k)g(2^{-m}) = g(r)$ and the corresponding terms under the right hand summations in (57) and (59) are equal (by (58)). Inequality in (57) follows from dropping terms (all positive) beyond $j = n - s$. Thus (56) and (57) hold, and by Theorem 17, these imply (48). Now suppose (48) holds, or equivalently (56) and (57) hold. We will prove (58) by induction on m.

If $m = 2$ then (56) gives $(r = \frac{1}{2})$

$$(61) \qquad \frac{3}{4}\, g(\tfrac{3}{4}) + \frac{1}{2}\, g(\tfrac{1}{2}) = \frac{1}{2}[g(\tfrac{1}{4}) + g(\tfrac{1}{2})] \quad \text{or} \quad g(\tfrac{3}{4}) = \frac{2}{3}\, g(\tfrac{1}{4})$$

which establishes (58) for $m = 2$. Now suppose (58) holds for $m = 2, 3, \ldots, m_0 - 1$. We will show it holds also for $m = m_0$. For any $k = 3, 5, \ldots, 2^{m_0} - 1$, let $r = (k-1)2^{-m_0}$. Then $r \in C_s$ where $s < m_0$. By the induction hypothesis the summations of (56) are equal term by term for $s + j < n$, that is for all terms except the last. Hence by (56) the last terms are equal. That is, $(r+2^{-m})g(r+2^{-m}) = rg(r-2^{-m})$, or $2^{-m}kg(2^{-m}k) = 2^{-m}(k-1)g(2^{-m}(k-2))$. This proves (58). The proof of (59) involves substituting terms given by (58). □

A simple sufficient condition can be given for the condition (59) of Theorem 18. For h defined by (60) we have

$$(62) \qquad \frac{h(2^j k-1)}{h(k)} = \frac{k+1}{k+2} \cdot \frac{k+3}{k+4} \cdots \frac{2^j k-2}{2^j k-1} < 1.$$

Now suppose that

$$(63) \qquad g(2^{-m}) \geq \sum_{j=1}^{\infty} g(2^{-m-j}) \qquad m = 1, 2, \ldots.$$

Then (62) and (63) imply (59). Thus (58) and (63) are sufficient conditions for (48). A further simplification is the restriction

$$(64) \qquad g(2^{-m-1}) \leq \frac{1}{2} g(2^{-m}) \quad m = 1, 2, \ldots$$

which implies (63). This leads us to a particular solution obtained by taking equality in (64) for each m. Arbitrarily taking $g(\frac{1}{2}) = \frac{1}{2}$ we get $g(\frac{1}{4}) = \frac{1}{4}$. $g(\frac{3}{4}) = \frac{2}{3} \cdot \frac{1}{4}$, $g(\frac{1}{8}) = \frac{1}{8}$, $g(\frac{3}{8}) = \frac{2}{3} \cdot \frac{1}{8}$, $g(\frac{5}{8}) = \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{1}{8}$, $g(\frac{7}{8}) = \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \cdot \frac{1}{8}$, $g(\frac{1}{16}) = \frac{1}{16}$, etc.

7. Payoffs which encourage honesty in the limit.

Consider the case of an infinite sequence of choices. Let $\mathfrak{a} = \{a_k\}_{k=1}^{\infty}$ where $a_k = -1$ or $+1$ again represents the forecaster's choice at the kth stage. Let $H(\mathfrak{a}, p)$ be the expected payoff for the forecaster with respect to the probability p. We can assume $p = P(A)$ is the true probability to be appraised.

If D is the set defined by (33), then for $r \notin D$, $0 < r < 1$, r defines a unique binary expansion and hence a unique sequence $\mathfrak{a}$. This sequence is given by

$$(65) \qquad r_{\mathfrak{a}} = \frac{1}{2} + \sum_{j=1}^{\infty} 2^{-j-1} a_j$$

where $r = r_{\mathfrak{a}}$. For $r \in D$, there exists two sequences of choices: one approaching r from the right and the other approaching r from the left. We will say q "encourages honesty in the limit" if for all sequences $\mathfrak{a}$ and $\mathfrak{B}$ such that $p = r_{\mathfrak{a}}$

(66) $\quad H(\mathfrak{A}, p) \geq H(\mathfrak{B}, p).$

If strict inequality holds in (66) whenever $p \neq r_{\mathfrak{B}}$ then g "strictly encourages honesty in the limit." These definitions are analogous to conditions (1) and (2) of Chapter I.

If for each $x \in (0, 1)$ we choose a corresponding $\mathfrak{A}$ satisfying (65) and denote $H(\mathfrak{A}, p)$ by $H(x, p)$, then $H(p, p)$ will satisfy condition (52). Define F and G by (51). Then Theorem 15 applies. If (52) holds then since $(0, 1) - D$ is dense in $(0, 1)$ and since Theorem 15 implies H is continuous, the value of $H(p, p)$ is independent of the choice of $\mathfrak{A}$ satisfying (65). That is, the forecaster's maximum expectation is independent of whether his choices $\{r_k\}$ defined by (30) approach r from the right or left for any $r \in D$.

The functions F, G and H are clearly given by $\lim F_n$, $\lim G_n$, and $\lim H_n$ respectively on the domain $(0, 1) - D$. It can be seen that (58), (59) and the existence of $\lim H_n(r, p)$ imply (66) or (52). However, there are infinite procedures which "encourage honesty in the limit" and do not encourage honesty for each n. We consider this class of procedures in this section.

Define the functions P_n and Q_n on $r \in (0, 1)$ by

(67) $$P_n(r) = 2^n[H(r, r) - H_n(r+, r)]$$

and

(68) $$Q_n(r) = 2^n[H(r, r) - H_n(r-, r)].$$

We will show (Lemma 5) that if g encourages honesty at the $\underline{n\text{th}}$ stage then for all $r \in (0, 1)$

$$P_k(r) \leq P_{k-1}(r) \quad \text{for all } k = 2, 3, \ldots, n \tag{69}$$

and

$$Q_k(r) \leq Q_{k-1}(r) \quad \text{for all } k = 2, 3, \ldots, n. \tag{70}$$

The limiting conditions of (56) and (57) can be expressed by

$$g(r) \geq \sum_{j=1}^{\infty} g(r-2^{-s-j}) \quad \text{for all } r \in D_s \tag{71}$$

and $s = 1, 2, \ldots,$ and

$$\lim_{n\to\infty} [P_n(r) - Q_n(r)] = 0 \quad \text{a.e. } r \in (0, 1). \tag{72}$$

The main theorem of this section is Theorem 19 which states these limiting conditions are such that g encourages honesty in the limit. The proof will be delayed to the end of this section.

Theorem 19.

The payoff g encourages honesty in the limit iff (69), (71) and (72) hold.

(71) is also the condition that $G_n(x)$ be nondecreasing for each n. The necessity of this condition is given in Lemma 4.

Lemma 4.

If g encourages honesty in the limit then $G_n(x)$ is nondecreasing and $F_n(x)$ is nonincreasing for all n.

Proof:

If (66) holds, then $G(x)$ is nondecreasing, $F(x)$ is nonincreasing, and $H(x, x)$ is continuous, independently of whether right or left continuity for the definitions of $F_n(x)$ and $G_n(x)$ for $x \in D_n$. This follows from Theorem 15. But $G_n(r) = G(r_n)$ if right-continuity is chosen and $F_n(r) = F(r_n)$ if left-continuity is chosen. This implies G_n is nondecreasing and F_n is nonincreasing for all n. □

The condition (69) is the condition that for all $x \in (0, 1)$,

$$H(x, x) - H_n(x, x) \leq \frac{1}{2} [H(x, x) - H_{n-1}(x, x)]; \tag{73}$$

the difference between the limiting H and H_n is at least halved at each stage. Lemma 5 shows this condition is necessary. Let $G(1) = G(1-)$, which must be finite if g encourages honesty in the limit.

<u>Lemma 5</u>.

If g encourages honesty in the limit then $P_n(x) \leq P_{n-1}(x) \leq H(x, x) \leq G(1)$ for all $x \in (0, 1)$.

<u>Proof</u>:

If (66) (or (52)) holds then by Lemma 4, (71) holds. Thus if $\{r_n\}$ is the sequence corresponding to a given $r \in (0, 1)$ and if left-continuity is assumed for F_n and G_n then

$$\begin{aligned} r_n g(r_n) &\geq r_n \sum_{j=n+1}^{\infty} g(r_n - 2^{-j}) \\ &= H(r_n, r_n) - H_n(r_n, r_n) \\ &\geq H(r, r_n) - H_n(r_n, r_n) \\ &= \sum_{\substack{k\geq n+1 \\ r_k > r}} r_k g(r_k) + r_n \sum_{\substack{k \geq n+1 \\ r_k \leq r}} g(r_k). \end{aligned} \tag{74}$$

If $r_n \geq r$ then (74) gives

$$(75) \qquad r_n g(r_n) \geq \sum_{\substack{k \geq n+1 \\ r_k > r}} r_k g(r_k) + r \sum_{\substack{k \geq n+1 \\ r_k \leq r}} g(r_k).$$

If $r_n \leq r$ then by multiplying both sides of (74) by $r/r_n \geq 1$ we get

$$(76) \qquad r g(r_n) \geq \sum_{\substack{k \geq n+1 \\ r_k > r}} r_k g(r_k) + r \sum_{\substack{k \geq n+1 \\ r_k \leq r}} g(r_k).$$

(75) and (76) imply

$$(77) \qquad H_n(r, r) - H_{n-1}(r, r) \geq H(r, r) - H_n(r, r).$$

This implies

$$(78) \qquad 2(H(r, r) - H_n(r, r)) \leq H(r, r) - H_{n-1}(r, r).$$

(78) implies $P_n(x) \leq P_{n-1}(x)$. $P_1(x) \leq H(x, x)$ because $g(x)$ is nonnegative. $H(x, x) \leq G(1)$ is obtained from (54) which gives for all $x \in (0, 1)$,

$$(79) \qquad H(x, x) = G(1) - \int_x^1 G(x)dx. \ \square$$

Define

$$(80) \qquad P(x) = \lim_{n \to \infty} P_n(x) \ , \ x \in (0, 1)$$

$$(81) \qquad Q(x) = \lim_{n \to \infty} Q_n(x), \ x \in (0, 1).$$

By the Lebesgue convergence theorem,

$$\int_x^1 P(z)dz = \lim_{n\to\infty} \int_x^1 P_n(z)dz \tag{82}$$

$$\int_x^1 Q(z)dz = \lim_{n\to\infty} \int_x^1 Q_n(z)dz. \tag{83}$$

Lemma 6.

If g encourages honesty in the limit then

$$H(x, x) = \int_x^1 (P(z) - Q(z) - G(z))dz + G(1). \tag{84}$$

Proof:

$$H(x, x) = \lim_{n\to\infty} H_n(x, x)$$

$$= \lim_{n\to\infty} \sum_{y\in D_n\cap[x,1)} (H_n(y-,y)-H_n(y+,y)+H_n(y+,y) - H_n(y+\tfrac{1}{2^n}-,y+\tfrac{1}{2^n})) + \lim_{n\to\infty} H_n(1-,1) \tag{85}$$

$$= \lim_{n\to\infty} \sum_{y\in D_n\cap[x,1)} 2^{-n}(P_n(y)-Q_n(y)-G_n(y+)) + G(1-)$$

$$= \lim_{n\to\infty} (\int_x^1 P_n(z)dz - \int_x^1 Q_n(z)dz - \int_x^1 G_n(z)dz) + G(1-) \tag{86}$$

$$= \int_x^1 (P(z) - Q(z) - G(z))dz + G(1-).$$

The last equality is given by (82), (83), and $G_n(z) \leq G(z)$, $G_n(z) \to G(z)$ as $n \to \infty$. □

The necessity of (72) follows from Lemma 6: (79) and (84) imply $\int_x^1 P(z)dz = \int_x^1 Q(z)dz$ for all $x \in (0, 1)$ and hence $P(x) = Q(x)$ a.e. $x \in (0, 1)$. We state this as Lemma 7.

Lemma 7.

If g encourages honesty in the limit then (72) holds, i.e., $P(x) = Q(x)$ a.e. $x \in (0, 1)$.

Proof of Theorem 19:

We have already shown in Lemmas 4, 5, and 7 that conditions (69), (71), and (72) are necessary for g to encourage honesty in the limit. Therefore we need only prove sufficiency.

Suppose (69), (71), and (72) hold. Then, as in Lemma 3, (85) and (86) hold. By (71), G_n is a non-decreasing sequence of nondecreasing functions and hence

$$\lim_{n\to\infty} \int_x^1 G_n(z)dz = \int_x^1 G(z)dz \quad \text{for all} \quad 0 \leq x \leq 1.$$

Clearly by the definitions of $P_n(r)$ and $Q_n(r)$ given in (67) and (68), (69) implies (70). Thus, if (69) and (72) hold,

$$\lim_{n\to\infty} \int_x^1 (P_n(z) - Q_n(z))dz = 0.$$

Therefore G is nondecreasing on $0 < x < 1$ and (54) or the equivalent equation (53) holds. Theorem 16 implies (52) holds independently of whether right or left-continuity was used in defining $H(x, x) = \lim H_n(x, x)$. Therefore (66) holds. □

The conditions (69), (71) and (72) could be made less stringent in Theorem 19. All that was necessary in the proof was that

(87) $$\lim_{n\to\infty} \int_x^1 [P_n(z) - Q_n(z)]dz = 0 \quad \text{for all } x \in (0, 1)$$

and that (71) hold. $P_n(r) - Q_n(r)$ is zero if $r \notin D$, and is given by

(88) $$P_n(r) - Q_n(r) = r \sum_{j=s+1}^{\infty} g(r-2^{-j}) - \sum_{j=s+1}^{\infty} (r+2^{-j})g(r+2^{-j})$$

if $r \in D_s$. Hence the following theorem expresses (66) in terms of g. The proof follows from the Lebesgue bounded convergence theorem.

Theorem 20.

The payoff g encourages honesty in the limit iff (71) and (72) hold and $|P_n(z) - Q_n(z)|$ is uniformly bounded on (0, 1).

CHAPTER V

Miscellaneous Related Problems

1. Consensus functions.

The payoff functions studied in this paper were to be given to individual forecasters. Other problems arise if the payoffs are to be given to a group of forecasters. The forecasters may give different answers if they wish to increase the total payoff to the group than they would give if they were concerned only about their own individual payoff. The forecasters may do better as a group if they pooled their densities to come up with just one average density.

When there are several forecasters, we need a consensus function $Q(p_1, p_2, \ldots, p_n)$ which gives a density representing the group of forecasters. The value of Q might depend on the forecasters' estimated densities $p_1, p_2, \ldots, p_n$, on a payoff function f, on a vector λ representing the client's belief about the relative expertise of each forecaster, and perhaps on the client's own estimated density.

For simplification, suppose one of the appraised densities $p_1, p_2, \ldots, p_n$ is known to be the true density p. Suppose the client has the degree of belief λ_k about the event $p = p_k$. Also, suppose the client expects to gain $H(\hat{p}, p) = \langle f(\hat{p}), p \rangle$ if p is the true density and he assumes it is $\hat{p}$, where $H(p, p)$ is convex. The client may then wish to choose $\hat{p}$ to maximize the expectation of $H(\hat{p}, p)$ under the distribution $P(p = p_k) = \lambda_k$.

Since

(89) $$E_\lambda(H(\hat{p}, p)) = < f(\hat{p}), \Sigma\lambda_j p_j >$$

and since f encourages honesty, the client should choose $\hat{p} = \Sigma\lambda_j p_j$. Thus, the client's expected "score" is best if he lets $Q(p_1, p_2, \ldots, p_n) = \hat{p} = \Sigma\lambda_j p_j$.

Other reasonable definitions can be given for Q. Suppose the n forecasters are to be paid $\lambda_k f(p_k)$ for their estimates p_k, $k = 1, 2, \ldots, n$ where $\Sigma\lambda_k = 1$ and f encourages honesty. Then the total payoff to the group is $u = \Sigma\lambda_i f(p_i)$ and is an element of the convex set C defined by (28) in Chapter III. If u lies on the boundary of C, we can treat it as though it were in the range of f and take $Q(p_1, p_2, \ldots, p_n) \epsilon \mathcal{P}$ normal to C at u. However, u will not usually be on the boundary of C unless $p_1 = p_2 = \ldots = p_n$. In general, one can choose a point u^* on the boundary of C which is "close" to u in some sense and then take $Q(p_1, p_2, \ldots, p_n) \epsilon \mathcal{P}$ normal to u^*, if such exists. For example, one can take $u^* \epsilon$ bdry (C) such that $\|u - u^*\| = \inf_{v \,\epsilon\, \text{bdry}(C)} \|u - v\|$. If $u^* = f(p) \epsilon C$ and C were smooth then $Q(p_1, p_2, \ldots, p_n) = p$. The distance $\|f(p) - u\|$ can be used as a measure of the precision of $p_1, p_2, \ldots, p_n$. It is not necessary that $\|\cdot\|$ be the $\mathcal{L}_2$ norm.

Another means for determining u^* is to find α such that αu lies on the boundary of C and then take $u^* = \alpha u$. Again, define $Q(p_1, p_2, \ldots, p_n) \epsilon \mathcal{P}$ normal to C at u^*, if such exists. The difficulty with this definition is that the value of Q depends on any constant added to the payoff function.

Winkler (1968) discusses and compares a number of methods for dealing with the consensus problem and presents the results of an experiment involving these methods. In an earlier paper, Winkler (1967b) gives empirical evidence showing a consensus may perform better than any individual.

2. Other properties of payoffs.

Because there are several payoffs which encourage honesty, one can consider more restrictive properties when making a decision about which payoff function to use. In many cases, the logarithmic payoff is the only payoff satisfying these more stringent conditions. For example, when the dimension of the vector space of random variables is finite and at least 3, the logarithmic payoff is the only continuous strictly honesty encouraging payoff whose value at (p, ω) depends only on the value $p(\omega)$, as already stated in Chapter 1.

Many properties of payoff functions imply the condition of dependence only on $p(\omega)$. For example, as Winkler (1969) stated:

> Since the likelihood used in a Bayesian analysis depends on the event which actually occurs, say X_h, and not on any other events which do not occur, a scoring rule is consistent with the use of likelihoods or log likelihoods to evaluate assessors only if the scoring rule depends solely on r_h and not on any $r_i \neq r_h$ [where r_h is the assessed probability of X_h].

Another very desirable property of a payoff function $f(p)(\omega)$ is that it is monotone nondecreasing in $p(\omega)$. This is defined by the following condition:

(90) $$f(p)(\omega) \geq f(q)(\omega) \quad \text{if} \quad p(\omega) > q(\omega).$$

If this condition does not hold, then one appraiser could be paid more if $\omega \in \mathfrak{X}$ occurred than another appraiser who assigned a larger likelihood to ω. A score or payoff not satisfying (90) in the discrete case thus seems unfair. In Euclidean space R^n, this condition and continuity imply again the condition of depending only on $p(\omega)$, and hence imply the logarithmic payoff if f encourages honesty and $n \geq 3$. We state this as Theorem 21.

Theorem 21.

Let $\mathcal{P} = \{p: p = (p_1, p_2, \ldots, p_n), 0 \leq p_k \leq 1, \Sigma p_j = 1\}$. If f is continuous and monotone nondecreasing in $p(\omega)$ then $f(p)(\omega)$ is a function only of $p(\omega)$, where $p \in \mathcal{P}$ and $\omega \in \mathfrak{X}$.

Proof:

Suppose $p, q \in \mathcal{P}$ and $p(\omega) = q(\omega)$ for fixed $\omega \in \mathfrak{X}$. Let $p_1 \in \mathcal{P}$ and $p_1(\omega) \neq p(\omega)$. Without loss of generality, assume $p_1(\omega) > p(\omega)$. Then (90) implies $f(\lambda p_1 + (1-\lambda)q)(\omega) \geq f(p)(\omega)$ for all $0 < \lambda \leq 1$. By letting $\lambda \to 0$, continuity of f implies $f(q)(\omega) \geq f(p)(\omega)$. Similarly $f(p)(\omega) \geq f(q)(\omega)$. Thus $f(p)(\omega) = f(q)(\omega)$. □

Good (1970) suggests that an appropriate logarithmic payoff might be $A \log(q(\omega)/\pi(\omega)) + B$ where $q(\omega)$ is the forecaster's estimate of $p(\omega)$ and $\pi(\omega)$ is the client's estimate. One advantage in the continuous case of this payoff over the logarithm of the probability density is that the payoff is invariant under a transformation of the variable ω. Other properties of the payoff $A \log(q(\omega)/\pi(\omega)) + B$ are additivity over independent trials of an experiment, and a new concept termed splitativity by Good (1970).

Splitativity in the discrete case is closely related to invariance in the continuous case. A payoff is splitative if its value is unchanged whenever the kth event is split into two events by a known randomizing device. Good (1970) gives a class of functions which have the property of splitativity or invariance and of encouraging honesty. Good suggests the problem of choosing a payoff from this class which "maximizes the forecaster's incentive in some sense, but without increasing the 'expected' fee."

There are appealing properties which could be considered which are not shared by the logarithmic payoff, The example given above in Chapter 1, due to de Finetti and Savage (1969), where the loss $L(r, p) = H(p, p) - H(r, p)$ is a function only if $r - p$ is one such example. Fortunately, the class of honesty encouraging payoff functions is large enough to contain functions satisfying additional conditions.

3. Undominated classes of probability measures.

The subject of encouraging honesty could be further generalized by taking $\mathcal{P}$ to be any set of probability measures on a space $(\mathfrak{X}, \mathfrak{A})$, not necessarily dominated, and taking $\mathfrak{L}_1$ again to be the set, defined by (8), of random variables whose expectations are defined for every $P \in \mathcal{P}$.

Let f be a mapping of $\mathcal{P}$ into $\mathfrak{L}_1$. We can use the same conditions (1) or (2) as before for f to encourage honesty. Let

H be a mapping of $\mathcal{P}$ into R, and substitute $E(f(Q)|P)$ for the inner product in conditions (9), (10), and (11). We could make the following definition:

Definition 7.

$q^* \in \mathcal{L}_1$ is a subgradient of H at $Q \in \mathcal{P}$ if for all $p \in \mathcal{P}$,

$$H(P) \geq E(q^*|P) - E(q^*|Q) + H(Q). \tag{91}$$

In the present context, Theorem 7 is still true, and the proof is the same. The question of the existence of a subgradient q^* for each $Q \in \mathcal{P}$ again is a restriction on H. The class $\mathcal{P}$ no longer is a subset of $\mathcal{L}$.

4. Symmetrical sequential payoffs.

In Chapter III, the choice of payoffs at the kth stage was given by the choice between the random variables $(g(p_k), 0)$ with probability vector $(p, 1-p)$ or the random variable $(g(p_k), 0)$ with probability vector $(p_k, 1 - p_k)$. This choice was equivalent to the choice between the constant payoff

$$X_1 = (p_k g(p_k), p_k g(p_k)) \tag{92}$$

and the payoff

$$X_2 = (g(p_k), 0) \tag{93}$$

where the distribution of each was given by the probability vector $(p, 1 - p)$.

The reason for using the random payoffs given by (92) and (93) was that their expectations were equal only at the point $p = p_k$

and hence forced the forecaster to decide whether $p \geq p_k$ or $p \leq p_k$. There are other payoffs which have this property. They are in fact of the form

$$X_j = (a_j(1-p_k),\ (x-a_j)p_k) \quad j = 1 \text{ or } 2. \tag{94}$$

The particular payoff defined by $x = 0$ and $a_1 = -a_2 < 0$ is symmetrical:

$$X_1 = (a(1-p_k),\ -ap_k) \tag{95}$$

$$X_2 = (-a(1-p_k),\ ap_k). \tag{96}$$

We can consider a sequence as described in Chapter III, of choices defined by (95) and (96) rather than by (92) and (93). The value of a at the <u>kth</u> stage is allowed to depend on p_k. The solution to the problem of choosing the function a such that the payoffs encourage honesty hopefully will be given in a later paper.

As mentioned in Section 2 of Chapter I, the class of choices at each stage can consist of three or more variables. For example, the intervals $A(p) \subset (0, 1)$ can be sequentially divided by thirds rather than halved. The forecaster can be given his choice of the three random variables

$$X_1 = (a(1-p_k),\ -ap_k)$$
$$X_2 = (0,\ 0)$$
$$X_3 = (-a(1-p_k^*),\ ap_k^*),\ p_k^* = p_k + 3^{-k-1},\ a < 0$$

at the <u>kth</u> stage, where it is to be decided if $p < p_k$, $p_k \leq p \leq p_k^*$

or $p > p_k^*$. A general theory might be given about the properties of such honesty encouraging increments of honesty encouraging cumulated payoffs.

REFERENCES

Aczel, J. and Pfanzagl, J. (1965). Remarks on the measurement of subjective probability and information. Metrika 11 91-105.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. Monthly Weather Review 78 1-3.

Buehler, R. J. (1970). Measuring information and uncertainty. In Godambe (1970).

De Finetti, B. (1962). Does it make sense to speak of "good probability appraisers?" In Good (1962), 357-364.

De Finetti, B. and Savage, L. J. (1969). The elicitation of personal probabilities and expectations. Unpublished.

Godambe, V. P. (ed.)(1970). Proceedings of the Symposium on Foundations of Statistical Inference (at Waterloo, Ontario, Canada). Holt, Rinehart and Winston of Canada, Toronto.

Good, I. J. (1952). Rational decisions. J. Royal Stat. Soc. Ser. B. 14 357-364.

__________ (ed.)(1962). The Scientist Speculates. Basic Books, New York.

__________ (1970). Information, rewards, and quasi-utilities. Unpublished.

Halmos, P. R. (1957). Introduction to Hilbert Space. Chelsea Publishing Co., New York.

Jeffrey, R. C. (1965). The Logic of Decision. McGraw-Hill, New York.

Köthe, G. (trans. D. J. H. Garling)(1969). Topological Vector Spaces I. Springer-Verlag, Berlin.

McCarthy, J. (1956). Measures of the value of information. Proc. Nat. Acad. Sci. 42 654-655.

Marschak, J. (1960). Remarks on the economics of information. Contributions to Scientific Research in Management. Univ. of California Press, Berkeley.

Ramsey, F. P. (1965). The Foundations of Mathematics. Littlefield, Adams and Co., Totowa, New Jersey.

Rockafellar, R. T. (1970). Convex Analysis. Princeton University Press, Princeton, New Jersey.

Savage, L. J. (1954). The Foundations of Statistics. Wiley, New York.

Shuford, E. H., Jr., Albert, A., and Massengill, H. E. (1966). Admissible probability measurement procedures. Psychometrika 31 125-145.

Valentine, F. A. (1964). Convex Sets. McGraw-Hill, New York.

Winkler, Robert L. (1967a). The quantification of judgement: some methodological suggestions. J. Amer. Statist. Assoc. 62 1105-1120.

__________ (1967b). The quantification of judgement: some experimental results. Proceedings of the Business and Economic Statistics Section, American Statistical Association, 386-395.

__________ (1968). The Consensus of Subjective Probability Distributions. Management Science 15 B-61-75.

__________ (1969). Scoring Rules and the Evaluation of Probability Assessors. J. Amer. Statist. Assoc. 64 1073-1077.